PHOTOPEAK DETECTION AND QUANTIFICATION USING WAVELET ANALYSIS

BY

JIE LU

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Nuclear, Plasma and Radiological Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2013

Urbana, Illinois

Master's Committee:

Assistant Professor Clair J. Sullivan, Director of Research
Assistant Professor Yang Zhang

ABSTRACT

Automated isotope identification has long been an important problem in homeland security and nuclear emergency response. This process is difficult for low-resolution spectra because of the presence of the Compton continuum, electronics noise, and peak overlap. The wavelet transform stands out among many potential solutions of this problem, owing to its ability to de-noise noisy signals, pattern matching, and simultaneous multi-resolution signal analysis. In this thesis, a novel wavelet-based algorithm for detecting peaks and measuring their areas is introduced. Its abilities in locating peaks, resolving overlapping peaks, and determining peak areas are presented and assessed with both simulated signals and real gamma-ray spectra. Peak area uncertainty was explored and future work and directions were discussed at the end.

ACKNOWLEDGMENTS

First and foremost, I am grateful to my parents and family for their unfailing love, support, and for being the reason I seek knowledge.

I would like to express my sincere gratitude to Dr. Clair Sullivan for her great support and patience throughout the entire work. Her insight, guidance and encouragement have great impact on my study and life. In addition, I highly appreciate the help of Dr. Yang Zhang and Dr. James Stubbins, and thank them for their time and efforts to help me improve the thesis.

In addition, I would also like to thank Jacob Stinnett, Mark Kamuda, Gianluca Panici, Benjamin Sturm and all other members in Radioisotope Identification Group for their useful discussions and inputs.

I would also like to extend my appreciation to Dr. Hsingtzu Wu, Dr. Nan Li, Dr. Liang Cai, Xiaochun Han, Xiang Ni, Sapna Maun, Umer Hassan, Yijia Pan and all other friends for sharing the good times and the bad times with me. Finally, I sincerely thank Becky Meline, Gail Krueger, Idell Dollison and all other faculty and staff of NPRE for their helps and guidance.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

A	Area of peaks (scalar)
AUC	Area Under Curve
B	Basis matrix 2048×2048
b	b(ch) = Background 1×2048
C_k	Covariance matrix for k 2048×2048
C_S	Covariance matrix for S 2048×2048
ch	Channel number
DRF	Detector Response Function
f	Simulated Signal 1×2048
g	Gaussian 1×2048
I	Identity matrix 2048×2048
k	Area vector(LS result for S) 2048×1
k_j	The j^{th} element in area vector k
k^o	True value of k
m	Number of elements in S (scalar)
n	Number of elements in k (scalar)
ns	White noise 1×2048
S	Signal vector (Wavelet-transform coefficient) 2048×1
S^o	True value of Wavelet Transform of the spectrum
s_{max}	Maximum scale

s_{opt}	Optimal scale = $s_{opt(E)}$ 1×2048
$T(E, s)$	Wavelet coefficients 512×2048
WTMM	Wavelet Transform Modulus Maxima
σ_A	Standard deviation of area (scalar)
σ_n	Standard deviation of noise (scalar)
$\psi(t)$	Wavelet (kernel) function
$\hat{\psi(\omega)}$	Fourier transform of wavelet function $\psi(t)$
ω	Frequency

CHAPTER 1

INTRODUCTION

Spectral analysis is among the most important steps in radioisotope identification. Today the most widely used detector for nuclear emergency response and border examinations is sodium iodide (NaI) scintillators owing to the fact that they are inexpensive, efficient, reasonably stable over a broad range of environmental conditions (for example, they can operate at room temperature without cooling). Therefore, demand exists for radio-isotope identification algorithms for gamma ray spectra acquired by scintillators, however it is a complicated problem [2].

1.1 The Isotope Identification Problem

Since the energy resolution of NaI is considerably poor and and the complexity of gamma-ray interactions with materials is significant [3] [1], the spectrum generated by NaI detectors is not as desirable as high-resolution detectors like high-purity germaniums because the energy needed to generate charge carriers is greater in scintillators. That leads to bad resolution for peaks in the spectra. Because the full-width-at-half-maximum of individual peaks is broad, the centroids of peaks are hard to accurately determine. Additionally, the Compton continuum, peak overlap, and electronics noise can cause interference with peak localization. How to measure information about photopeaks such as the peak centroid, peak area, and uncertainty of the peak area is a major concern.

Attempts have been made for peak detection by using Gaussian functions [4]. The generalized second derivative of the Gaussian function is calculated and convolved with the spectrum to eliminate the interference from the fluctuating baseline and that from Compton edges. In the end, peak amplitude, centroid, and width are determined by least squares fitting. It was almost a primitive way of generating wavelet functions according to the input signal. The wavelet-equivalent function was found by optimizing the two determining parameters: the degree of smoothness and the support (region of definition) of the function, which is not in the scope of this thesis but the possible future work of this thesis.

The aforementioned method provides a direction for discriminating peak features from a fluctuating baseline and Compton continuum and demonstrated ability in handling overlapping double peaks, but peak areas are not solved from this method. Secondly, the peak width needs manual iteration to optimize. Efficient automation of peak searching and quantification needs to be found.

The weighted least squares method is also utilised in alpha spectroscopy to fit peaks with Gaussian functions [5]. Peak areas and uncertainties are determined in the study. However, the presence of Compton continuum and significant baseline problems do not exist in alpha spectral analysis. The least squares method demonstrated in the study here alone can not solve the problem in gamma-ray spectral analysis.

1.2 Our Solution to the Problem

In this thesis we describe a method to locate peaks and quantify their area based on wavelet analysis. Similar problems in mass spectrometry were also studied and used wavelet analysis well [6]. In mass spectrometry, the x-axis would be mass-to-charge ratio and the y-axis is the intensity of materials

with corresponding mass to charge ratio. The goal is to qualitatively or quantitatvely evaluate the peaks of interests and determine the presence of certain materials and the mass of them. Peak detection is critically important in mass spectral analysis and the ability of continuous wavelet transform is demonstrated in peak detection. With the help of the wavelet transform modulus maxima (related concepts will be introduced in Chapter 2) lines, peaks were successfully located in this study. However, peak areas were not further pursued here.

Wavelet analysis is good at extracting singularities or features of interests [7] [8]. For quantification of these features, Non-negative Least Squares(NNLS), in our case, can be used. NNLS has displayed satisfactory ability in peak quantification not only for related applications but also for other kinds of spectral analysis. For example, full spectrum analysis (FSA) for environmental in-situ gamma-ray spectra measured by a NaI detector was studied with Non-Negative Least Squares method and confirmed with results from an HPGe detector with small error [9]. In this method, NNLS is used for fitting the whole spectrum with a linear combination of background and the fundamental spectra profiles of each library isotope. The criterion for this study was the achievement of minimum chi-square value. However the limitation is also obvious. In-situ gamma-ray spectrum analysis has predefined knowledge of isotopes that would be present in the soil samples which contains a few isotopes, mainly K-40, U-238, Th-232, Cs-137. Also the true values of peak area and the uncertainty were not presented in the discussion.

Instead of Gaussian functions used in the FSA method, exponential kernel functions were exploited as basis function template in the kinetic modeling method to decompose the spectrum for the purpose of quantifying peaks [10]. After the decomposition, the correlation matrices and covariance matrices can be obtained by NNLS calculation and these matrices are critical in uncertainty evaluation of the decomposition and quantification. Both kinds of matrices are investigated respectively considering both the coupling of model parameters and the pre-known noise distribution.

Another demonstration of NNLS in spectral analysis is Nuclear Magnetic

Resonance (NMR). As a non-destructive diagnostic technique, it offers abundant industrial and medical applications. Among them are quality control and pharmaceutical metabolism monitoring with respect to moisture measurements by observing signals produced after electromagnetic wave interactions with water and aqueous solutions [11]. During the process, individual nuclear magnetic moments interact with the environment (lattice) but also each other. The reaction between nuclear magnetic moments and the lattice is called spin-lattice relaxation, and is characterized by decay time constant T1. The reaction between nuclear magnetic moments themselves is referred to as spin-spin relaxation, characterized by T2, the spin-spin relaxation decay time [12] [13]. The intensity value on y-axis reveals the number of proton on frequency value(got from relaxation time T2 value) indicated on x-axis. Because of this property, NMR spectrum reveals the concentrations of components in the samples. With the quantification process of NNLS algorithm, the concentrations of components are calculated with respect to their corresponding relaxation time for given NMR spectrum. In this application, the basis kernel function is exponential due to the nature of decay time. By minimizing the least-squares variance, NNLS finds the best fitting spectrum in a least squares sense. Even though the fit of models to the spectrum is not unique, requiring further statistical techniques, NNLS performs well in spectrum quantification.

After all being said, the spectroscopic analysis problem is eventually boiled down to peak detection and quantification. Direct fitting of peaks manually is cumbersome. With the wavelet transform, peak detection can be done much more efficiently and automatically. For peak quantification linear regression methods have come a long way, as described in diverse area such as mass spectrometry, nuclear magnetic resonance data analysis, molecular spectroscopy [6] [14] [15] [16]. Therefore, combining the ability of pattern recognition of the wavelet transform and the quantification capability of least squares fitting methods, a solution for gamma-ray spectrum analysis is proposed in this work. The solution is composed of two parts. The first part is peak detection achieved by wavelet analysis, following by the second part– peak quantification with Non-negative Least Squares method (NNLS). Elaboration on the fundamental theories and implementation is given in chapters

below.

CHAPTER 2

THEORY

2.1 Wavelet Basics

A wavelet function is a continuous function with finite support. Rather than fluctuating on and on like sinusoids, it only fluctuates in its support region. Because of this property, it possesses a much better capability in capturing the location of certain singularities while convolving with the input signals [8] [17] [18]. Compared to the Fourier transform which is often seen in signal processing applications, the wavelet transform is a powerful tool to project signals simultaneously in both the time domain (because we are dealing with energy spectrum instead of time series signal in our case, the time domain here means energy domain) and frequency domain (equivalent to the scale domain) [7]. The Fourier transform can only allow us to process signals in the time domain or in the frequency domain one at a time. This empowers us to have a thorough understanding on both the position and the frequency of specific features of interests. The coefficients produced after the transform reveal where and how greatly wavelet function matches the special features in the original signals. Normally features not of interest would be regarded as noise, which would be eliminated with the proper choice of wavelet function. Therefore, the wavelet transform basically combines the processes of de-noising and template fitting into one single step. Because of this distinguished pattern recognition capability, wavelet analysis is widely recognized in image processing [8] [19] [7], climate signal detection [20], defect detection in mechanical engineering [21], upstream exploration analysis for well logging [22], financial engineering [23] and proteomic pattern recognition [24].

Mathematically, the continuous wavelet transform is written as below:

$$T(E,s) = \int_{-\infty}^{\infty} \psi(\frac{\mathrm{t}-\mathrm{E}}{\mathrm{s}})\mathrm{f(t)}dt; \tag{2.1}$$

where T is the wavelet transform coefficient, $f(t)$ is the original signal, $\psi(t)$ is the mother wavelet kernel function, E is the shifted distance in the daughter wavelet function and s is the scale of the daughter wavelet function $\psi(\frac{t-E}{s})$. $\psi(t)$ can not claim itself a wavelet function unless the *admissibility conditions* are satisfied [25] [7] [8]:

(1) $\psi(t)$ is square integrable, which means it has finite energy:

$$\int |\psi(\mathrm{t})|dt < \infty \& \int |\psi(\mathrm{t})|^2 dt < \infty; \tag{2.2}$$

(2) the Fourier transform of $\psi(t)$: $\hat{\psi}(\omega)$ satisfies:

$$\int \frac{\left|\hat{\psi}(\omega)\right|^2}{\omega} d\omega < \infty; \tag{2.3}$$

which requires that $\hat{\psi}(0) = 0$.

(3)

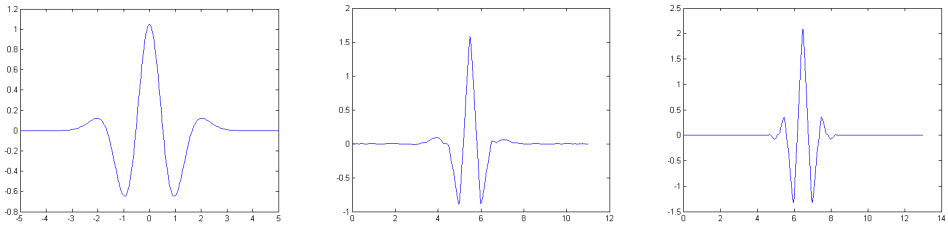
$$\int \psi(\mathrm{t})dt = 0; \tag{2.4}$$

which is derived from (2), meaning no energy contributes to the zero-frequency component (i.e. the DC energy is zero).

With the admissibility conditions satisfied, the continuous wavelet transform (CWT) is defined as the convolution of a signal with the daughter wavelet functions $\psi(\frac{t-E}{s})$, which are the scaled and shifted versions of mother wavelet function. The result is that the one dimensional original signal, $f(t)$, is projected on a two dimensional vector space with independent variables time t and scale s.

There are countless numbers of series of wavelet kernel functions for applications: Haar, Morlet, Daubechies [26], and so on. The features of interest of this study would be photopeaks in gamma-ray spectra which have approximately Gaussian profiles. With that in consideration, theoretically we are choosing near symmetric wavelets in order to pick up the Gaussian pro-

file and obtain the corresponding scales of the peak [17]. Mother wavelet functions such as: 'gaus4'–the fourth derivative of a Gaussian, 'coif2'–coiflet wavelet of second order smoothness (differentiable up to two degrees) and 'bior2.6'–biorthogonal wavelet of second order smoothness in reconstruction function and sixth order smoothness in decomposition function, were chosen for evaluation based on previous study [18]. Their profiles are shown in Figure 2.1. They have finite range and among the three only 'gaus4' has an analytic expression.



(a) gaus4 **(b)** coif2 **(c)** bior2.6

Figure 2.1: Profiles of some wavelet templates

Three Gaussian functions were generated in the upper part of Figure 2.2, with peak centroids located at 140, 560, 900 and σ of 5, 10, and 20 respectively. The wavelet transform was performed on all of them using 'bior2.6' wavelet. A two-dimensional matrix of wavelet transform coefficients is formed from the process, which is called the scalogram, as shown in Figure 2.2. The local maxima in the scalogram are called wavelet transform modulus maxima (WTMM) and are shown as black lines in the figure. By orderly collecting and linking these points, WTMM lines are formed.

In previous studies it was shown that the larger coefficients from the CWT mean better matching between signal and wavelet [6]. The paper concluded peak features can be found on WTMM lines which are formed by linking local maxima of CWT coefficients across the scales. The SNR threshold specifically defined by CWT coefficients instead of the simple peak-amplitude-to-noise-level is claimed to have a decreased rate of false positive alarms on peak detection. The advantage of this method is that it is possible to avoid the trouble of baseline removal and smoothing. Even though it is not pursued

further to get the AUC (area under curve) in the spectrum, measurement of the AUC obtained from the wavelet method is pointed out as a future research to this thesis. As will be shown later, WTMM lines were used to identify useful information in scalogram: the maximum scale , and the wavelet transform coefficients of the signal on that scale, which were used in the next step (NNLS calculation) to solve the accurate value of peak centroid and peak area.

For peak features, if the wavelet transform coefficient values are plotted with respect to scales along the line, a profile with the slope which changes from positive to negative number is expected to show up, as shown in the lower part of Figure 2.3. In this profile, the scale where largest wavelet coefficient in y axis can be found is defined as the maximum scale s_{max}. What the maximum scale means is at this scale the wavelet function resonates the most with original signal. The maximum scales for the three Gaussian functions shown in Figure 2.2 are 39, 79, and 162.

As is seen in the three Gaussian peaks example, s_{max} is positively correlated to the width of peak features. Therefore, as the energy resolution decreases with energy and the lower limit of peak width increases in real spectra, s_{max} actually increases with energy as well. If the the relationship of peak's FWHM (full width at half maximum) versus the peak centroid in units of channel for certain detector is acquired, the relationship of maximum scale s_{max} with the corresponding channel can be predicted according to that. It is defined in this work as the detector response function (DRF). Take the DRF of a commercial Ortec 2×2 NaI detector for example: Figure 2.4a presents the relation between peak centroid and peak FWHM. According to our study, this relationship for a NaI detector is best conditioned when fitted with a polynomial function in the energy domain. Since the wavelet transform is a linear transform, the DRF in the wavelet domain remains as a polynomial function. The curve in Figure 2.4b gives the relation between peak centroid and optimal scale s_{opt} which indicates the scale of single photopeak on that channel if any of single photopeak exists.

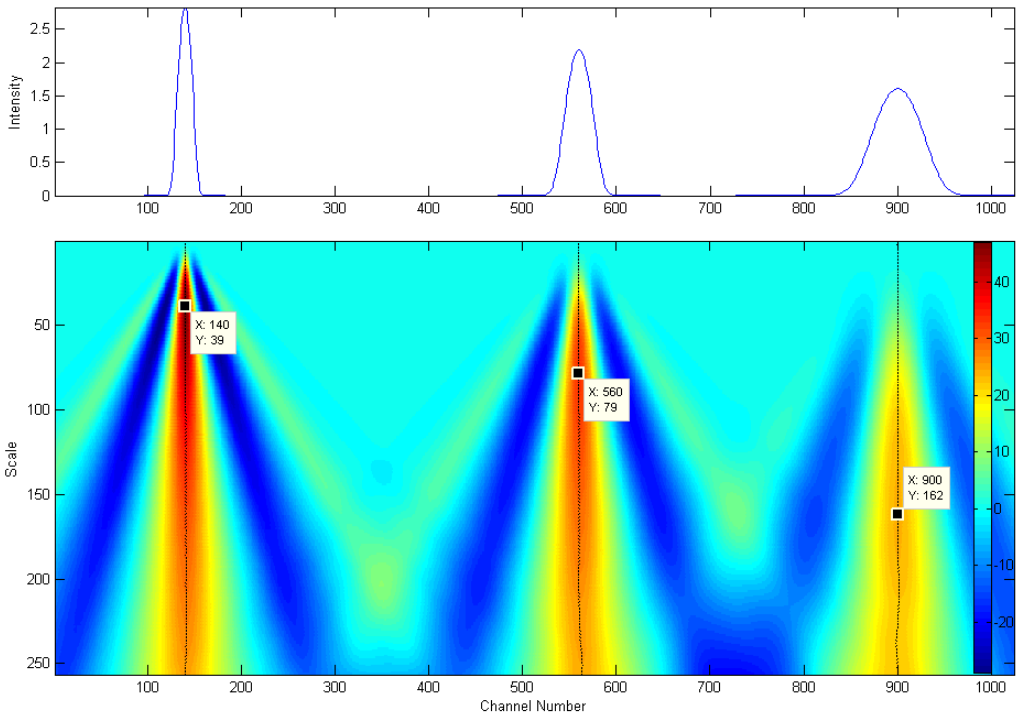


Figure 2.2: Gaussian functions and their wavelet transform scalograms, with maximum scales shown.

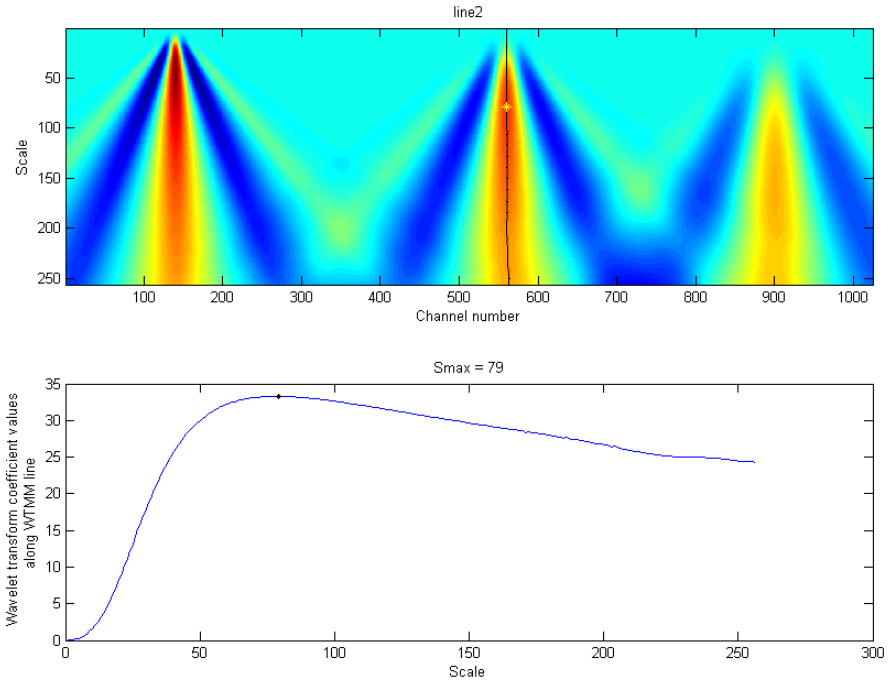
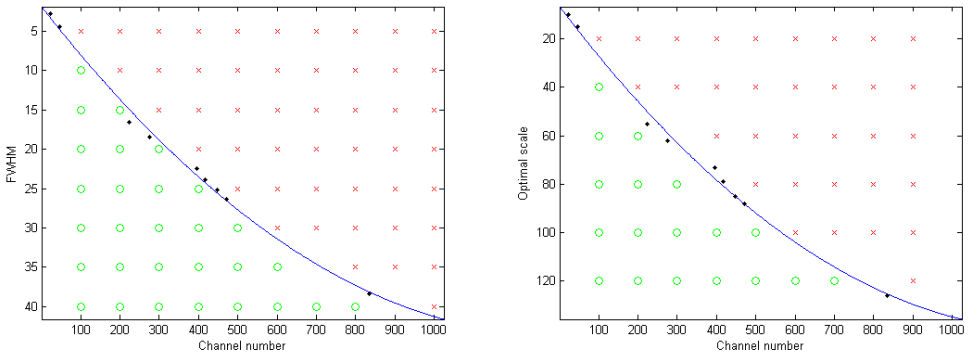


Figure 2.3: WTMM line for the middle peak and the wavelet transform coefficient values along the line.

As we can see in Figure 2.4b, the DRF in wavelet domain drawing a line between possible region and impossible region for true peaks to show up, represents the optimal scale s_{opt} for the given detector. The maximum detected scale s_{max} of a true peak measured with this detector on certain channel can not be smaller than the optimal scale s_{opt} for same channel on the DRF curve. This is because s_{opt} is the lowest limit of width of single true peak that could be provided by gamma photon. The width can not go any smaller. Any narrower peaks would be caused by electronics noise fluctuations. So any wavelet transform modulus maxima located on a scale in the region marked with red cross indicates a feature with narrower width than a true peak's should be. However in situations where $s_{max} > s_{opt}$, there might be multiple peaks presented or other effects that broaden the peaks.



(a) DRF in energy domain **(b)** DRF in wavelet domain

Figure 2.4: Profiles of DRF functions and the separated regions: green circles indicate possible region for true peaks because $s_{max} \geq s_{opt}$, whereas red crosses show the region where $s_{max} < s_{opt}$ indicating no true peak is present.

2.2 Non-Negative Least Square Fundamentals

Non-negative least squares (NNLS) is a special case of the least squares problems with linear inequality constraints. NNLS offers a potential advantage for

gamma-ray spectral analysis because all photopeaks will have non-negative areas. By imposing this extra non-negative constraint in solving the linear regression problem by least squares, theoretically the peak area should be acquired with least error, or best goodness of fitting [27] [28].

In a nut shell, the algorithm aims to solve the problem $S = Bk$ with minimum $\|Bk - S\|$ (minimum error in the least square sense), where S and B are known and k in turn remains to be solved. On top of this basic statement of the problem, inequality constraint of $k \geq 0$ is also imposed as additional requirement, making it an NNLS problem. The algorithm solving this problem is named NNLS algorithm. Basically k is assumed as many times as needed to achieve Bk as close to S as possible. To this point, based on the linear property of wavelet transform, the area of peak is proportional to the value of k. Mathematically, it is written as

$$\int \mathrm{kf(t)}dt = k \int \mathrm{f(t)}dt. \tag{2.5}$$

There are two main ingredients needed in the algorithm: a vector matrix, S(in our case S is a vector of wavelet transform coefficients of the original spectrum at the correct scale), and matrix, B, which is known as the designed matrix or the basis matrix [29] (in our case B is a square matrix which possesses our pre-known knowledge of the scale of the peaks). Other ingredients are defined or presumed in the following process of the algorithm. Detailed steps of the algorithm are described in following box [27] [28].

Steps:

1. initialize inputs: m×n basis matrix B, m×1 original signal vector S, empty set P (defined as empty n-vector in Matlab), set Z=$\{1, 2, ..., n\}$, integer iter = 0 and empty n-vector solution k

2. define dual vector $w = B^{\intercal}(S - Bk)$ and check terminating conditions: all elements in the final k are ≥ 0 ; if satisfied, break and return k and norm of the dual vector $\|w\|$

3. if Z is empty and all elements in the final dual vector w reach ≤ 0

4. find the index t $\in Z$ that $w(t)$= max$\{w_j$: j $\in Z\}$ and move t to P from Z

5. set B_P = the t^{th} columns of basis matrix B where t $\in P$

6. solve for z so that B_P z $\cong S$ (z is current dummy for k in this loop)

7. if $z_j > 0$ for all j $\in P$ then k = z and go to step 2

8. iter = iter + 1, terminate if number of iterations of this optimization exceeds the limit number $3n$, let k = z and return k and norm of the dual vector $\|w\|$; find the index q $\in P$ that $\frac{k_q}{k_q - z_q}$ = min $\{$ $\frac{k_j}{k_j - z_j}$: $z_j \geq 0$, j$\in P$ $\}$

9. set $\alpha = \frac{k_q}{k_q - z_q}$

10. set k = k + α(z - k)

11. find all indices j that $k_j = 0$ and j $\in P$ and move them from P to Z and go to step 6

12. end and return k and norm of the dual vector $\|w\|$

CHAPTER 3

IMPLEMENTATION

In this chapter, the wavelet-based algorithm is described. The algorithm will be explained with a Cs-137 spectrum and consists of two main parts. The first part is the wavelet analysis where the centroids of the peaks are detected. The second part is NNLS calculation where the exact values of peak centroid and peak area are determined. Matlab is used to generate the code for all analysis and calculations. The code is provided in Appendix.

3.1 Wavelet Analysis

The goal of wavelet analysis is to detect the presence of a true peak. If the peaks detected are true peaks, their full widths at half maximum (FWHMs) should satisfy the energy resolution limits calibrated for given detector. As indicated in Chapter 2, the FWHM in energy domain is positively correlated to scale in wavelet domain, hence, this problem can be transfer to the wavelet domain and the maximum scale of a true peak should satisfy the limit calibrated by the DRF in wavelet domain. The maximum scale s_{max} of a peak is found where the peak resonates the most with the selected wavelet function. In order to find this maximum scale s_{max}, the largest coefficient is pursued for its location in the scalogram.

3.1.1 FILTER 1: WTMM-DRF crossing filter

The search for s_{max} can be performed on every WTMM lines, but it is time-consuming and not necessary. Because a lot of WTMM lines are created by noise, baseline fluctuations, or the Compton continuum, they do not cross the DRF curve and are discarded. Filtering out these WTMM lines can improves the overall algorithm performance. Therefore, the first filter in the wavelet analysis process is the WTMM-DRF crossing filter. In Figure 3.1, the wavelet transform of Cs-137 spectrum is performed with the 'bior2.6' wavelet function. WTMM lines (black solid lines) formed by true photo-peaks always cross with DRF curve measured by commercial Ortec 2×2 NaI detector with a 1024 channel multichannel analyzer(white solid line). Figure 3.2 shows the filtered WTMM lines that remain in the peak detection process.

3.1.2 FILTER 2: Vertical filter

The second filter is the so-called "vertical filter" on the remaining WTMM lines [18] [17]. WTMM lines that deviate over too large a range with respect to energy are eliminated and are not considered for the next step. In the current Cs-137 spectrum, WTMM lines shown in red in Figure 3.3 would be regarded as too diagonal for the spectrum and are eliminated from further consideration.

3.1.3 FILTER 3: Profile check on curve of wavelet transform coefficient values

The third filter in wavelet analysis is the profile check on the curve of wavelet transform coefficients along wavelet transform modulus maxima lines. For true peaks, curves of their wavelet transform coefficients along wavelet trans-

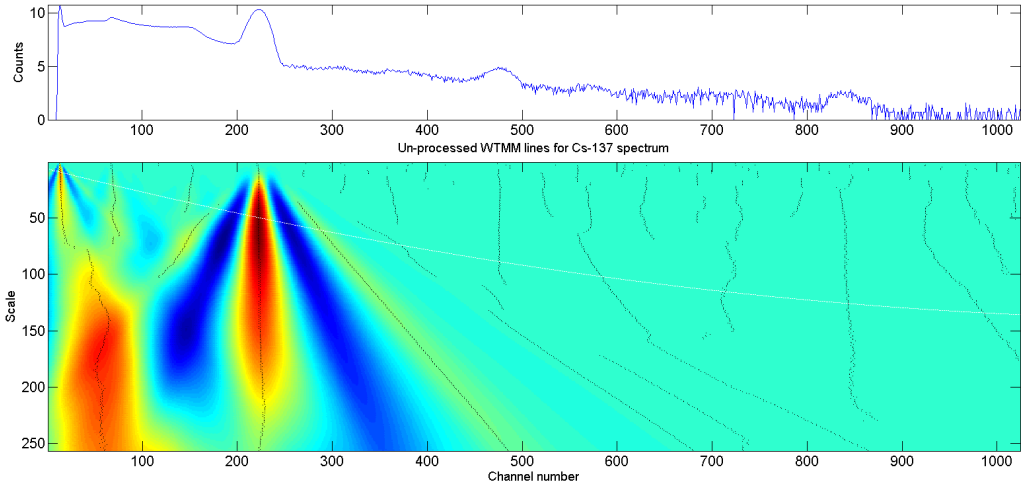


Figure 3.1: The un-processed WTMM lines for Cs-137 spectrum

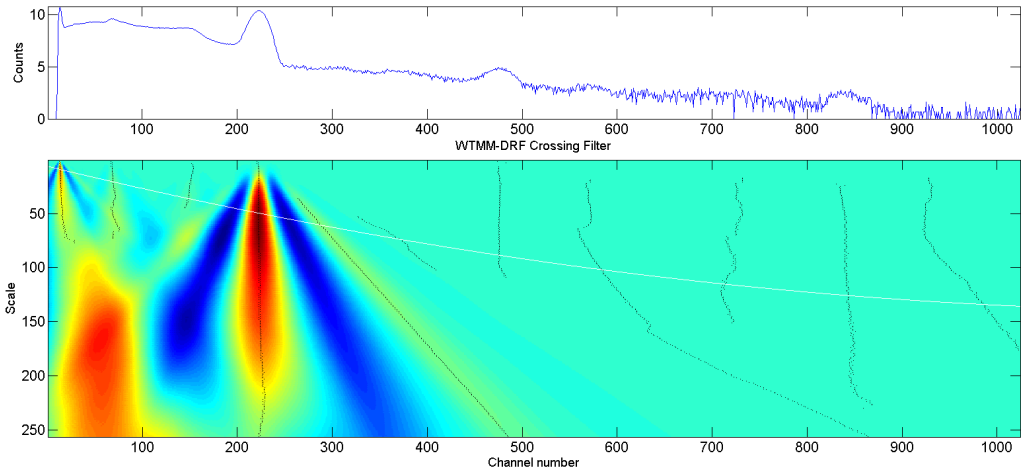


Figure 3.2: The WTMM lines remaining for Cs-137 spectrum after the first filter

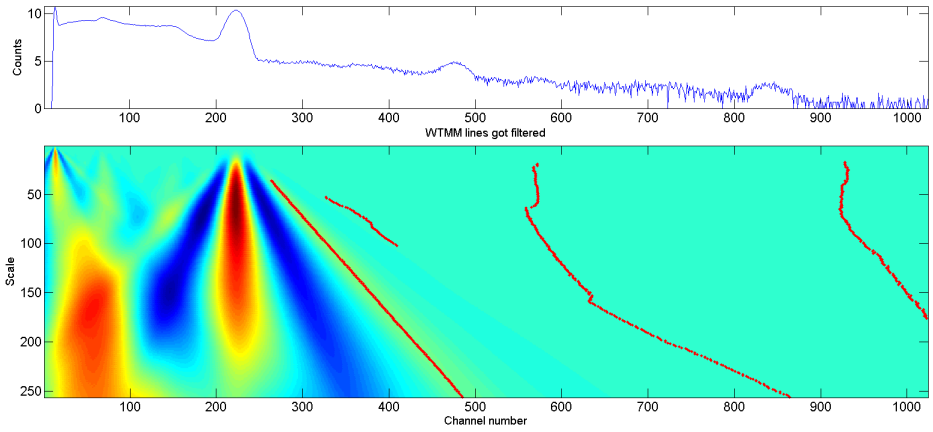


Figure 3.3: The WTMM lines eliminated after the second filter

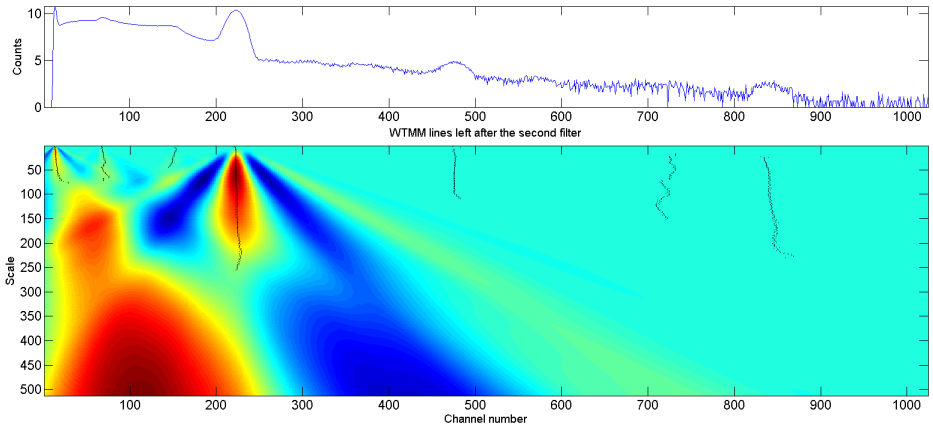


Figure 3.4: The WTMM lines remaining after the second filter

form modulus maxima lines possess a sign change of the first derivative of the curves themselves. The blue line in lower part of Figure 3.5 shows the profile of a true peak. In the Cs-137 spectrum all the WTMM lines from the second filter possess profile like this one. For the current version of the code, the s_{max} is the first local maxima in the WTMM line profile.

3.1.4 FILTER 4: $s_{max} \geq s_{opt}$

As discussed in Chapter 2, s_{max} of true peaks from the spectrum are greater than or equal to s_{opt} on that channel. Therefore, the fourth filter obeying this criterion is imposed on the remaining WTMM lines to get rid of electronic noise. Peaks with WTMM lines that pass this filter are potential candidates and would be sent to last filter with scale comparison on DRF function before zero-padding preparation for the NNLS calculation.

As for peak quantification with NNLS calculation following wavelet analysis, the optimal scale acquired from the wavelet transform modulus maxima needs to be used to generate not only the basis function matrix but also the DRF (detector response function) if simulated spectra with noise and baseline or real spectra are dealt with.

There is certain pit fall during wavelet transform process on real spectra though. Unlike noiseless and baseline-free signals where both the left and the right sides are continuously zero and there is little boundary effect in the transform, real spectra or simulated signal with noise and baseline demand the elimination of boundary effect.The presence of non-zero baseline through and beyond the net peak region, Compton scattering continnum could be captured by the convolution with the wavelet functions and trigger extra solutions outside the peak region. For this reason, zero-padding was executed [30]. Detail on this will be discussed below in the next section on NNLS calculation.

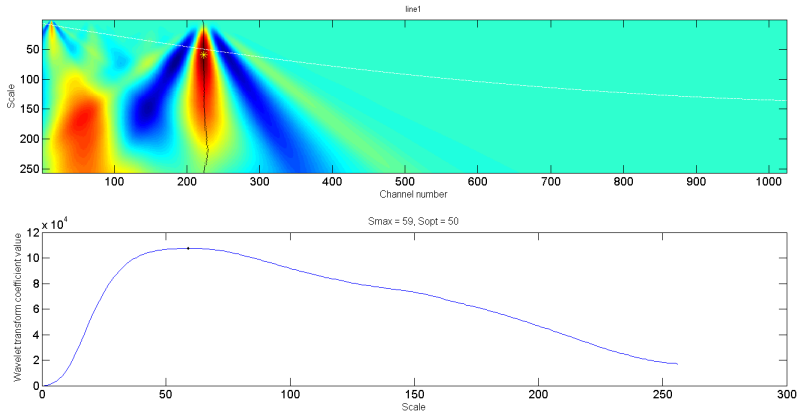


Figure 3.5: WTMM line with its WT coefficient value profile

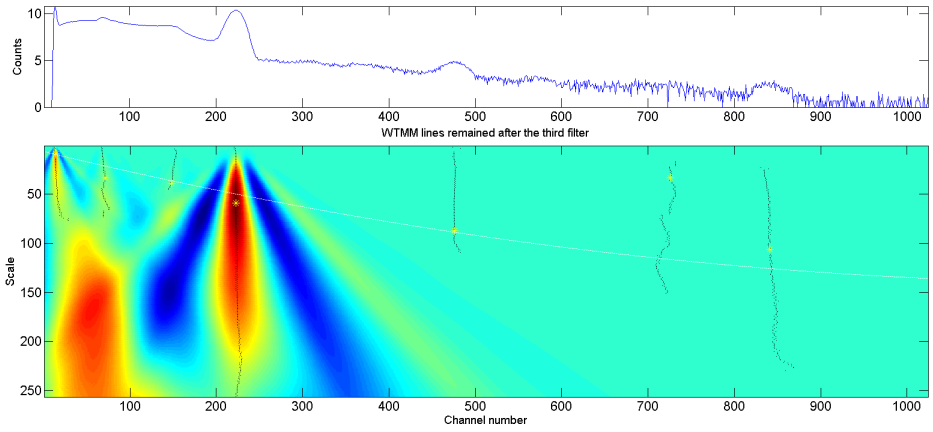


Figure 3.6: WTMM lines after the third filter

3.2 NNLS Solution

After peaks are identified from wavelet analysis, their optimal scale wavelet transform coefficients are used for NNLS calculation. The requirement of non-negative peak areas is established. As discussed in the last section, to eliminate boundary effects, the region within $\pm\ 3\sigma$ from the peak centroid would be cut out and padded with zeros by both sides before NNLS calcu-

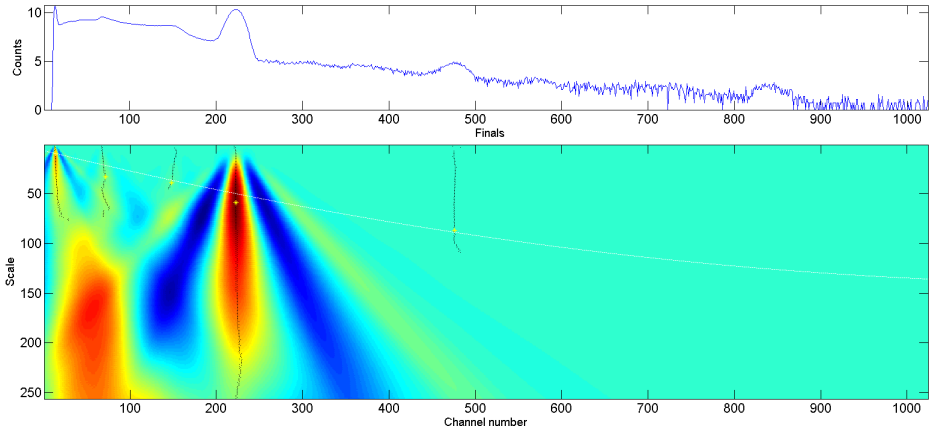


Figure 3.7: WTMM lines survived after the fourth filter

lation is performed. The *lsqnonneg* function is called in Matlab to generate the NNLS solution vector, k, and norm of residuals.

CHAPTER 4

RESULTS ON PEAK CENTROID AND PEAK AREA

4.1 Simulated Spectra

Before the combined method of wavelet analysis and NNLS calculation was tried on real spectra, simulated signals were used in testing the method. Signals of single peaks with or without white noise and baseline continua were tried to evaluate the method's ability in eliminating the interference of features other than peaks. Additionally, signals of multiple equally-scaled or differently-scaled peaks with or without white noise and baseline situations were tried to evaluate the method's ability to resolve overlapping.

4.1.1 Single peak detection

As shown in Figure 4.1, a single Gaussian peak with area of 50 and centroid of channel 903 was simulated as shown by the blue solid line. After wavelet analysis and NNLS calculation, the centroid was found at channel 903 and the area was measured to be 50, as shown by the green line of the same figure.

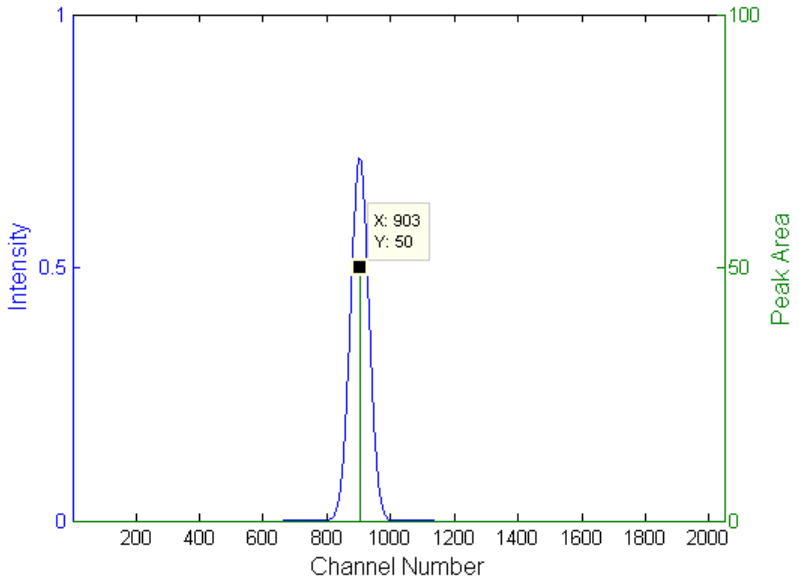


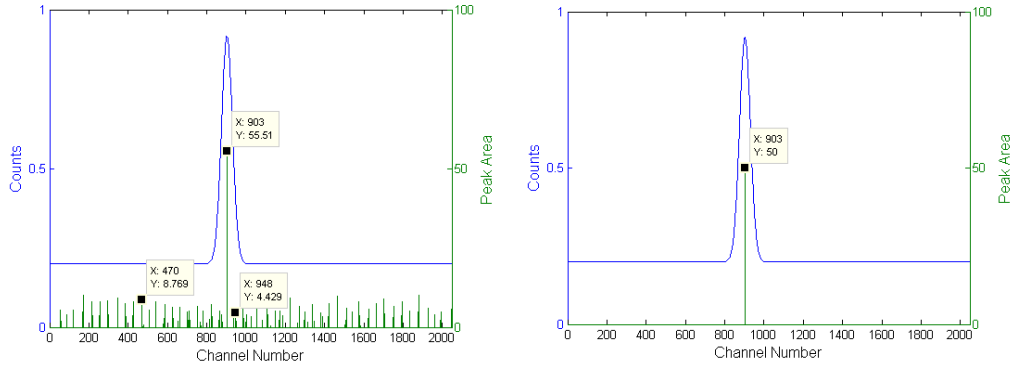
Figure 4.1: Simulated single peak signal with centroid at 903 and area of 50

Peak area evaluation with an added linear baseline

A constant baseline was added to simulate situations where a peak is situated on a continuum caused by Compton scattering and background radiation. As shown in Figure 4.2a, multiple spikes of solution values were given by NNLS calculation but some of them are not even in the peak region. This was caused by boundary effect from area extended beyond the net peak region. It was then corrected by applying zero padding, with results presented in Figure 4.2b. The outcome is that the patched wavelet/NNLS algorithm was able to restore 100% accuracy for centroid and peak area when a constant baseline is presented.

In Figure 4.3, a signal was simulated combining a linear baseline and a Gaussian peak. An extreme situation is simulated where the baseline was so large that the Gaussian peak was not immediately visible. However, the wavelets were able to detect the peak and its area was solvable as shown in Figure 4.4 on page 25. In this case where an overwhelming linear baseline was added, the algorithm can still obtain accurate results on peak centroid, but the calculated peak area value was (49.99+0.005347)=49.995347, a deviation from

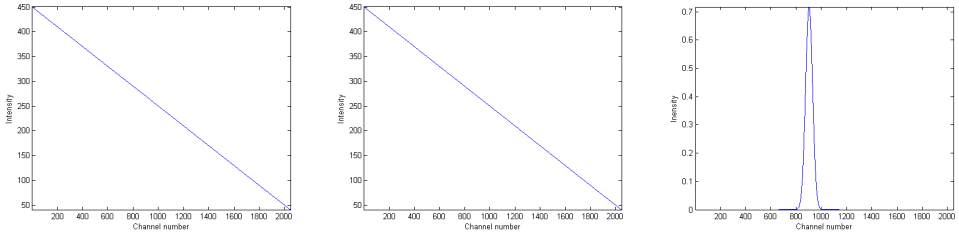
the true value 50 by around 1.33%.



(a) Single peak with baseline not processed with padding

(b) Single peak with baseline after processed with padding

Figure 4.2: Boundary effect elimination in simulated peak added with horizontal baseline (without noise).



(a) s(ch) **(b)** y(ch) **(c)** g(ch)

Figure 4.3: Combined signal $s(ch) = y(ch) + g(ch)$, consisting of a Gaussian peak, $g(ch)$, with area 50 at channel 903 and a linear baseline given by $y(ch) = -0.2 \times ch + 0.2$, where ch is channel number.

Deviation of peak area evaluation when added with baseline and white noise

The importance of uncertainty evaluation comes mainly with noise where the residual of the NNLS solution increased with the magnitude of the noise and the peak area started to deviate from the true value. In the following Table 4.1, each value of area is the average of 10 simulations and the variance

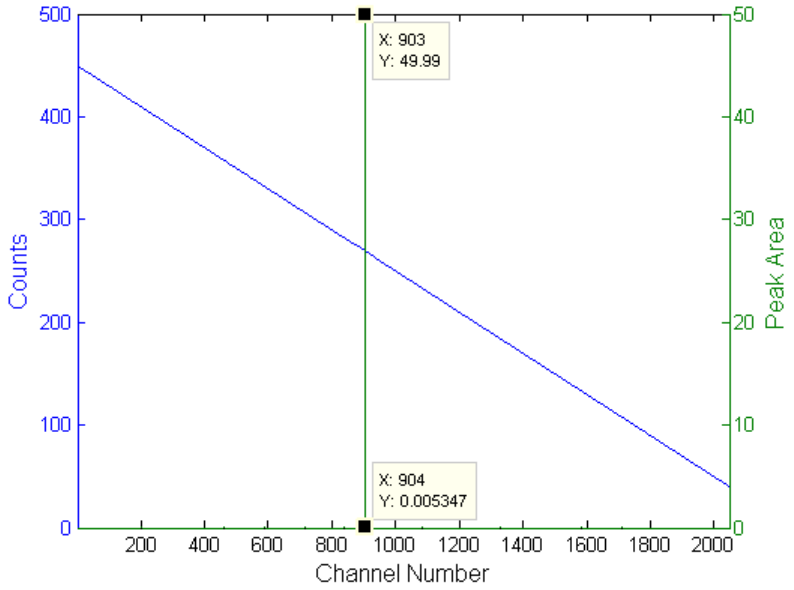


Figure 4.4: Solution for peak area in simulated tilted-baseline-added signal. Spikes in peak region are summed to yield the value of area.

is obtained from each batch of 10 simulations for every noise level. Some examples of these signals are shown in Figure 4.5. The conclusion is when the magnitude fluctuation of the noise is so big that it reaches 0.03, even wavelet analysis combined with NNLS won't be able to accurately determine the peak area within a 2% deviation.

4.1.2 Multiple peaks situations

Centroids and areas for peaks with equal areas

One of the main obstacles in spectral analysis is resolving overlapping peaks. In this section, double peaks with identical width and area were simulated to assess the algorithm's resolving capability.

Table 4.1: Peak area deviation depends on the level of noise magnitude

Noise Level	0.001	0.005	0.01	0.02	0.03	0.04	0.05	0.1	0.2	0.3	0.4
Area Average	49.96	50.11	50	50.91	52	51.48	53.12	54.4	57.93	67.91	76.04
Variance	0.03	0.11	0.92	0.84	1.06	3.66	13.61	8.74	71.01	136.42	242.74

Table 4.2: Examination on double peaks with same area

Distance	Standard	10	11	12	13	14	15	16
Left peak area & Right peak area	50	50.21	51.1	50	51.33	49.32	49.43	50
	50	50.19	48.9	50	48.67	50.68	50.57	50
Left peak area deviation & Right peak area deviation	0%	0.4200%	2.2000%	0.0000%	2.6600%	-1.3600%	-1.1400%	0.0000%
	0%	0.3800%	-2.2000%	0.0000%	-2.6600%	1.3600%	1.1400%	0.0000%
Distance	Standard	17	18	19	20	21	22	23
Left peak area & Right peak area	50	50.48	50	50	49.664	50	50	50
	50	49.525	50	50	50.34	50	50	50
Left peak area deviation & Right peak area deviation	0%	0.9600%	0.0000%	0.0000%	-0.6720%	0.0000%	0.0000%	0.0000%
	0%	-0.9500%	0.0000%	0.0000%	0.6800%	0.0000%	0.0000%	0.0000%

As shown in Figure 4.6, double peaks with equal area 50 were simulated to assess the ability in accurately determining the centroids and areas of both the peaks. Peaks with equal area remain resolvable to each other in Figure 4.6 even when not visibly apparent to the naked eye. It was then obvious to examine the minimum distance between two peaks which have the same width and same area in the spectrum. Table 4.2 shows the result of this examination. All peaks in this part possess FWHM of 65.355 channels. The distance between peaks with equal area was decreased to 16 channels until the deviation reached larger than 1% and was considered not accurately resolvable any more. The conclusion is that as peaks get closer to each other, it is harder for the algorithm to accurately resolve their areas. The lowest boundary for distance between the two peaks is 0.245 FWHM of the peaks with same FWHMs.

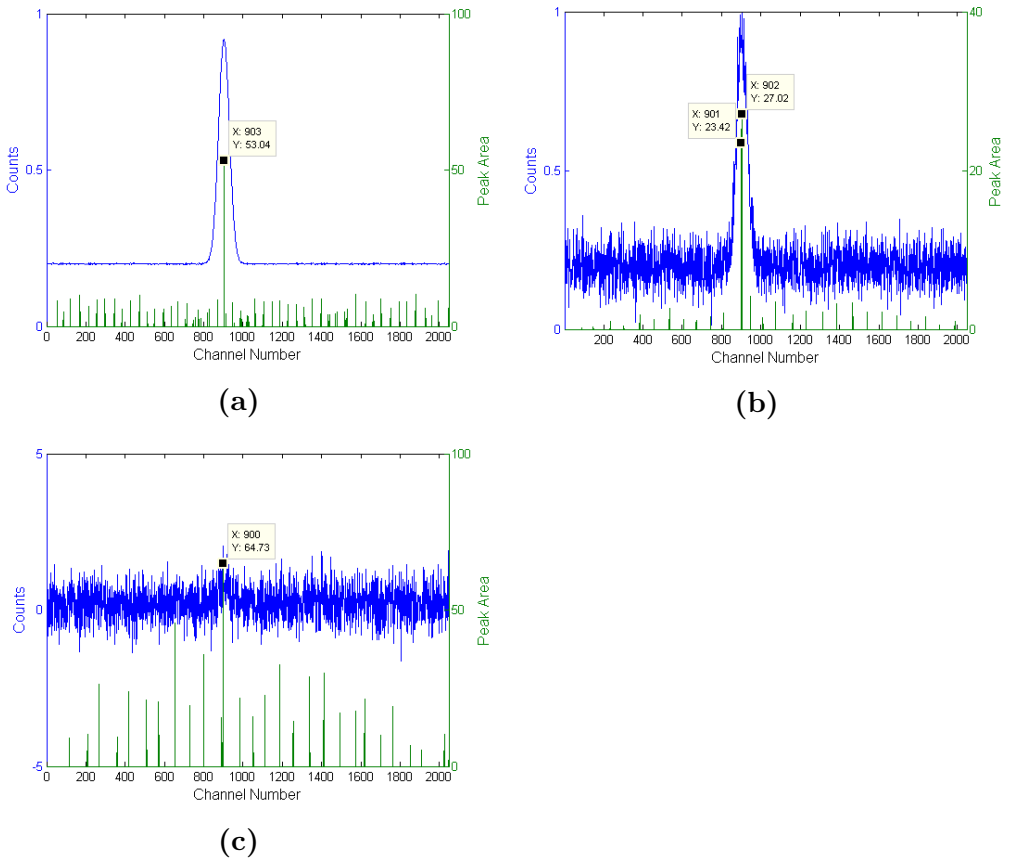


Figure 4.5: Examples of simulated single peak with baseline and white noise of different level: (a) Simulated signal with baseline and white Gaussian noise of σ 0.001 (b) Simulated signals with baseline and white Gaussian noise of σ 0.005 (c) Simulated signals with baseline and white Gaussian noise of σ 0.5

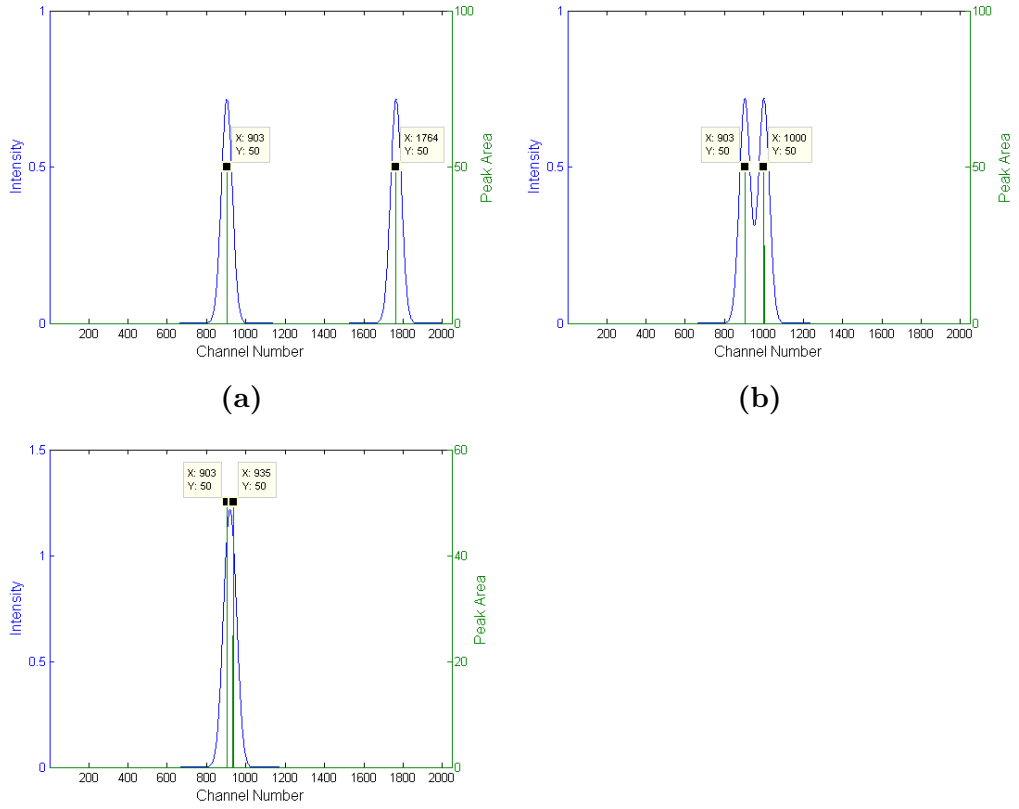


Figure 4.6: Multiple peaks situation:(a)Simulated double peaks with equal area 50 at channel 903 and 1764 (b)Simulated double peaks with equal area 50 at channel 903 and 1000 (c) Simulated double peaks with equal area 50 at channel 903 and 935

Centroids and areas for peaks with unequal areas

Real life situations do not always provide close peaks with equal area. More oftenly peaks with unequal area are present on the spectrum. Hence we evaluated the limit of minimum distance between two peaks that allows clear resolvability. Table 4.3 records the data from this evaluation. Using a limit on the deviation of 1%, the limits of minimum distance for 80%, 60%, 40%, 20% and 10% were 18, 22, 22, 23 and 23 channels respectively (All peaks are generated with FWHM of 65.355 channel). The limit of minimum resolvable distance increases when the area of the smaller peak decreases. A few examples are illustrated in Figure 4.7.

In summary, peak locations and areas were validated in this section 4.1 by analysing simulated signals with noise fluctuations, linear baselines and overlapping effects. The algorithm can accurately locate single peak centroids and determine single peak areas with either constant baseline or linear baseline. On top of that, single peak areas for signals added with white Gaussian noise calculated by the algorithm was kept in 2% deviation from true values while the magnitude fluctuation of the noise hits 0.03. For double peaks situation, resolvability on overlapping peaks' areas were assessed for this algorithm. The algorithm can resolve peaks with equal areas which are as close to only 16 channels apart while their FWHMs are both 65 channels. Double peaks with different areas are also resolvable by the algorithm but the limit of minimum distance between the two peaks increases with the difference between their areas. For area ratio of small peak area over large peak area being 80%, 60%, 40%, 20%, 10%, the minimum distances that allows to resolve the peaks are 18, 22, 22, 23, and 23 channels respectively while both peaks are generated with FWHM of 65 channels.

Distance (channel)	standard	10	11	12	13	14	15	16	17	18	19	20
Areas	50	48.995	50	50.35	50.96	46.84	49.5671	50.54	50.45	50	50	49.727
(right peak area/left peak area = 80%)	40	41.01	40	39.66	39.04	41.17	40.44	39.466	39.553	40	40	40.28
Deviation	0.0000%	-2.0100%	0.0000%	0.7000%	1.9200%	-6.3200%	-0.8658%	1.0800%	0.9000%	0.0000%	0.0000%	-0.5460%
	0.0000%	2.5250%	0.0000%	-0.8500%	-2.4000%	2.9250%	1.1000%	-1.3350%	-1.1175%	0.0000%	0.0000%	0.7000%
Areas	50	50	50.85	50	49.212	50.63	50.58	50.49	50	50	49.772	49.8082
(right peak area/left peak area = 60%)	30	30	31.144	30	30.798	29.3785	29.419	29.5134	30	30	30.23	30.19
Deviation	0.0000%	0.0000%	1.7000%	0.0000%	-1.5760%	1.2600%	1.1600%	0.9800%	0.0000%	0.0000%	-0.4560%	-0.3836%
	0.0000%	0.0000%	3.8133%	0.0000%	2.6600%	-2.0717%	-1.9367%	-1.6220%	0.0000%	0.0000%	0.7667%	0.6333%
Areas	50	54.47	54.5	49.88	50	50.63	50.4	50.51	50.31	49.6753	49.8553	50
(right peak area/left peak area = 40%)	20	15.533	15.5124	20.121	20	19.3785	19.604	19.488	19.692	20.328	20.143	20
Deviation	0.0000%	8.9400%	9.0000%	-0.2400%	0.0000%	1.2600%	0.8000%	1.0200%	0.6200%	-0.6494%	-0.2894%	0.0000%
	0.0000%	-22.3350%	-22.4380%	0.6050%	0.0000%	-3.1075%	-1.9800%	-2.5600%	-1.5400%	1.6400%	0.7150%	0.0000%
Areas	50	52.74	50	49.601	51.82	51.47	51.27	49.6602	49.8148	50.4	50.37	50.29
(right peak area/left peak area = 20%)	10	7.268	10	10.401	8.187	8.5342	8.733	10.346	10.186	9.599	9.64	9.713
Deviation	0.0000%	5.4800%	0.0000%	-0.7980%	3.6400%	2.9400%	2.5400%	-0.6796%	-0.3704%	0.8000%	0.7400%	0.5800%
	0.0000%	-27.3200%	0.0000%	4.0100%	-18.1300%	-14.6580%	-12.6700%	3.4600%	1.8600%	-4.0100%	-3.6000%	-2.8700%
Areas	50	51.33	51.14	51.62	51.4	51.23	50.55	50.47	50	50	49.9586	49.9684
(right peak area/left peak area = 10%)	5	3.565	3.86771	3.3834	3.6055	3.779	4.4535	4.5284	5	5	5.038	5.034
Deviation	0.0000%	2.6600%	2.2800%	3.2400%	2.8000%	2.4600%	1.1000%	0.9400%	0.0000%	0.0000%	-0.0828%	-0.0632%
	0.0000%	-28.7000%	-22.6458%	-32.3320%	-27.8900%	-24.4200%	-10.9300%	-9.4320%	0.0000%	0.0000%	0.7600%	0.6800%

Table 4.3: Double peaks with unequal areas

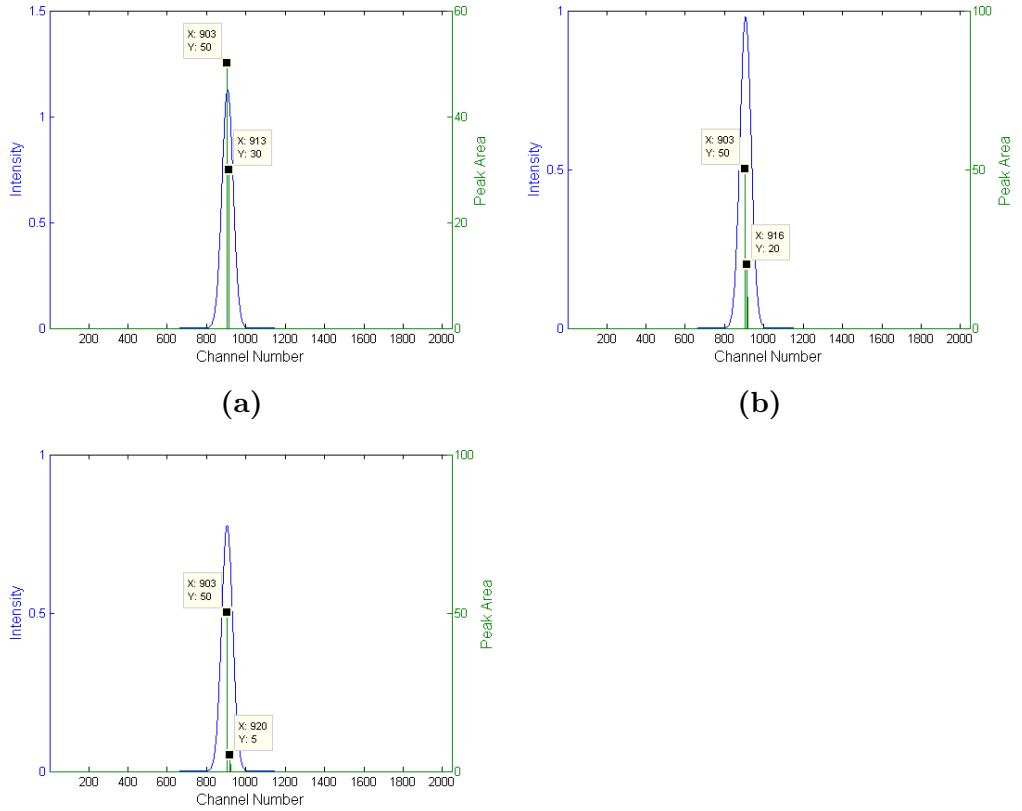


Figure 4.7: Double peaks with unequal area: (a) Simulated double peaks at a distance of 10 channels with area 50 at channel 903 and area 30 at channel 913 (b) Simulated double peaks at a distance of 13 with area 50 at channel 903 and area 20 at channel 923 (c) Simulated double peaks at a distance of 17 with area 50 at channel 903 and area 5 at channel 920

4.2 Real Spectra

4.2.1 The results from wavelet analysis for Am-241, Co-57, Cs-137, Mn-54, Co-60, Na-22, Background(K-40, Tl208)

Based on pre-measured data from Ortec model #2M/2 2×2 NaI detector with 1024 channel MCA, the DRF function and basis matrix were generated. Peak measurements extracted by the wavelet analysis are summarized below.

Isotope : Am-241

The predominant gamma ray of Am-241 spectrum is located at 59.5 keV. It is detected by the algorithm. The centroid of the peak determined by the algorithm is on channel 24, in consistent with the result from Origin. As for the peak area, the algorithm gave 1466709. Assuming Origin result 315787 was true value, the algorithm deviated by 364.461%. This deviation of peak area came from the fact that this 59.5 keV peak sitted on a much broader and higher peak. Future simulation should be carried out by simulating overlapping peaks with largely different FWHMs. and peeling out the impact of peak not of interest. Additionally, the algorithm detected the peak at 1460 keV emitted by K-40 in the background radiation but not the background peak of 2614 keV by Tl-208, as shown in Figure 4.8. Investigation of the Tl-208 peak revealed that there are multiple maxima on that WTMM line, as shown in Figure 4.9. Even though the centroid was correctly measured at channel 837, the correct s_{max} scale was not measured. The same case occurred for Tl-208 peaks in other real spectra too. Improvement on filters and DRF curve could be made for future implementations.

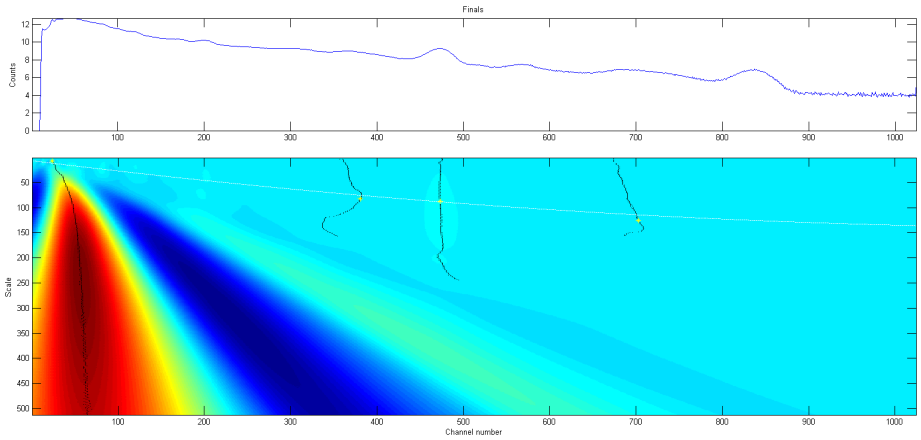


Figure 4.8: Features in the Am-241 spectrum identified.

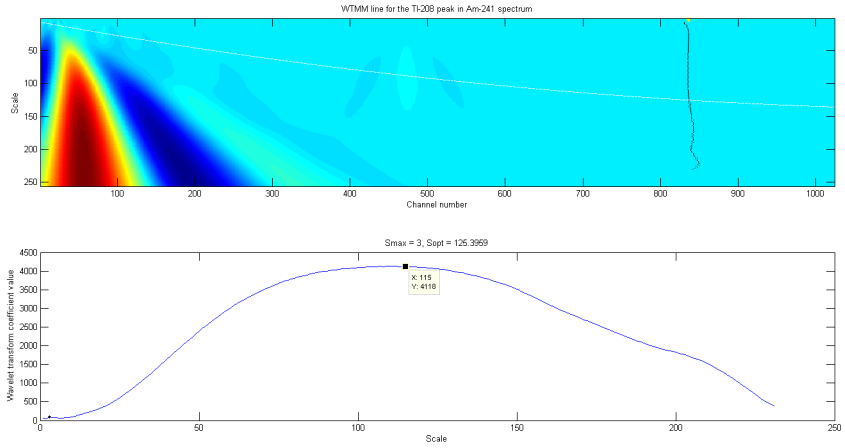


Figure 4.9: The Tl-208 peak not detected in spectra.

Isotope : Co-57

In the case of the Co-57 spectrum, the peak of 122 keV was detected at channel 46 as shown on Figure 4.10. The peak centroid for this 122 keV photopeak given by the algorithm is 46, in comply with result from Origin. The peak area was 642414 by the algorithm, having a 42.77% deviation when compared with 449963 by Origin. The deviation was caused by the peak of interest sitting on a non-linear baseline which assembles part of a broad peak. The K-40 peak of 1460 keV is missed due to the same reason as Tl-208 in Am-241 spectrum: the detected value of s_{max} was at a local maximum significantly smaller than s_{opt} versus the global max, as shown in Figure 4.11. The Tl-208 peak is missed in this spectrum too, but is a different situation from the K-40 peak. The maximum coefficient along the line occurs earlier than optimal scale. There is a 6% deviation to the optimal scale in DRF function, hence, the concern of carefully monitored measurement environment for DRF function is brought up here.

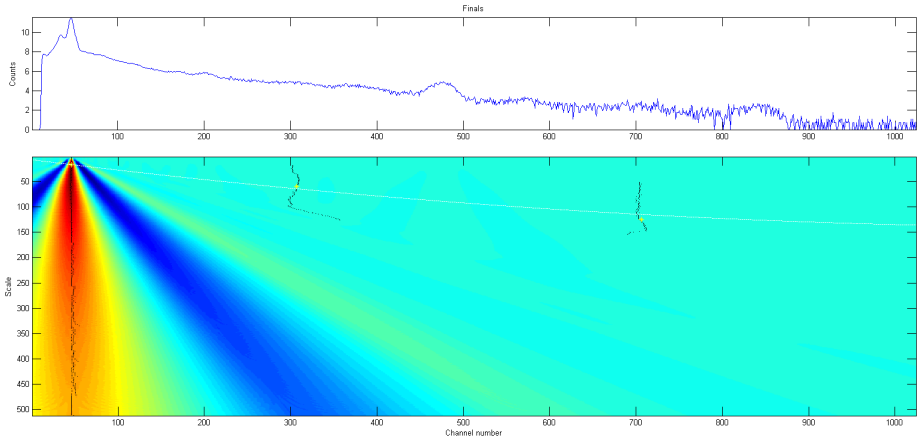


Figure 4.10: Features in the Co-57 spectrum that were identified.

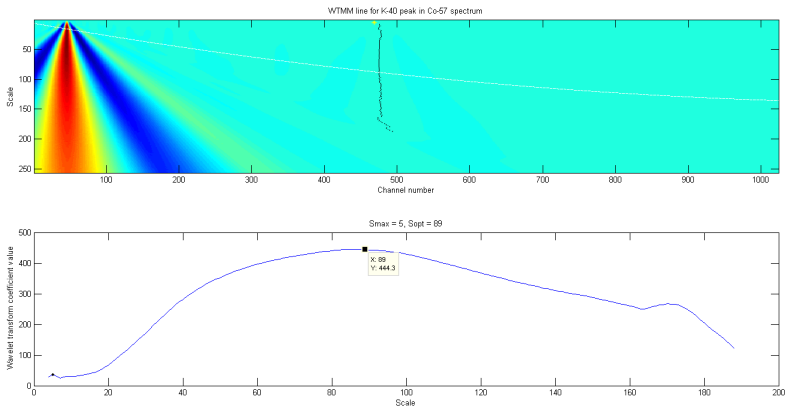


Figure 4.11: The WTMM line for the K-40 peak not identified in the Co-57 spectrum.

Isotope : Cs-137

The analysis of Cs-137 was successful, as shown in Figure 4.12. The backscattering peak of approximately 186 keV was obtained from the spectrum, together with the characteristic peak of 662 keV and background peak K-40 at 1460 keV. The Tl-208 peak at 2614 keV was missed for the same reason stated in Am-241 spectrum. The characteristic peak of 662 keV has peak centroid on channel 223. The peak area was calculated to be 508599 by the algorithm, deviating by 6.244%, when compared with 542473 determined by Origin.

Isotope : Mn-54

Figure 4.13 shows that characteristic peak of 835 keV in Mn-54 was detected. The peak centroid was found using the wavelet algorithm with perfect accuracy at channel 275 where true value of centroid was assumed by result from Origin. Other features like Compton edges were also detected by the wavelet

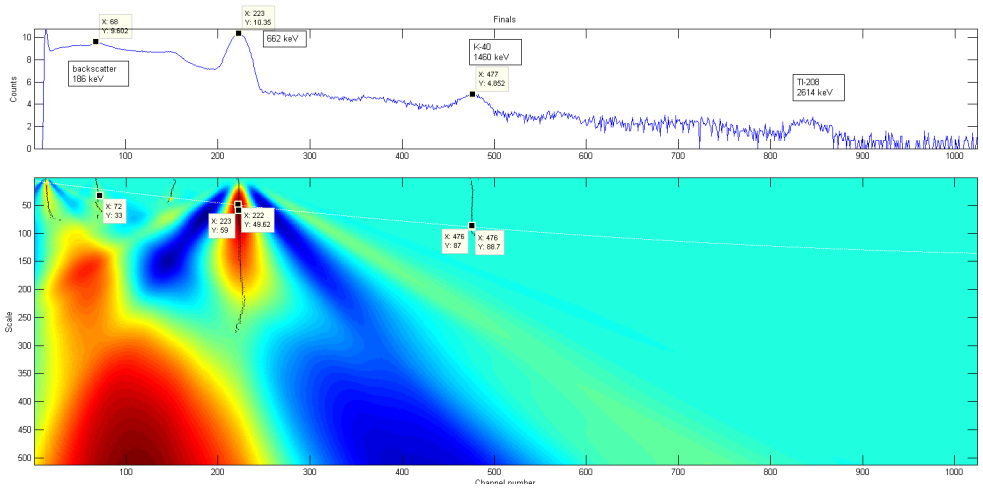


Figure 4.12: Features in the Cs-137 spectrum that were identified.

analysis algorithm. The peaks for K-40 and Tl-208 were missed in the same situation of the Tl-208 peak in the Am-241 spectrum: s_{max} measured much smaller than s_{opt} (Figure 4.14 and Figure 4.15). The peak area for 835 keV peak was 4771418 by the algorithm, having only a 1.225% deviation from Origin result 4830570.

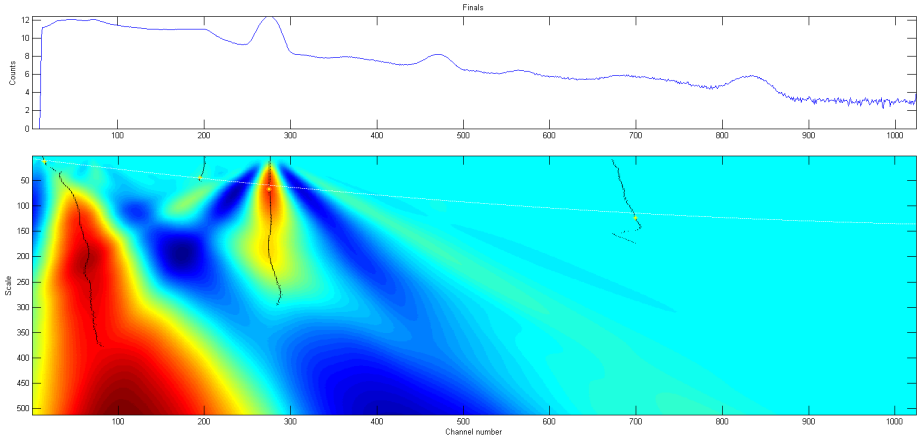


Figure 4.13: Features in the Mn-54 spectrum that were identified.

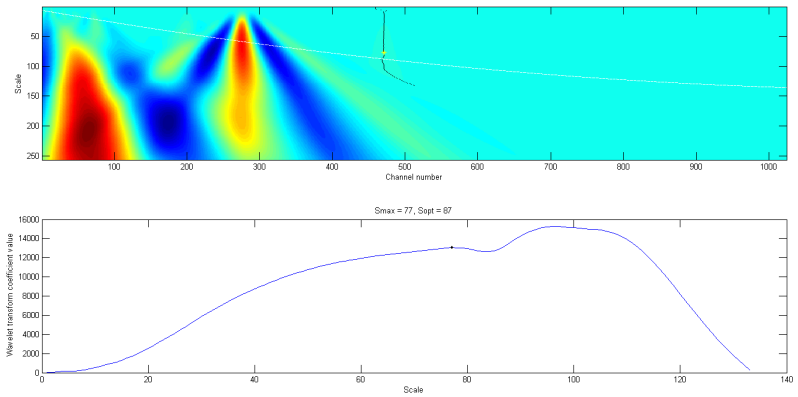


Figure 4.14: The missing K-40 peak in the Mn-54 spectrum.

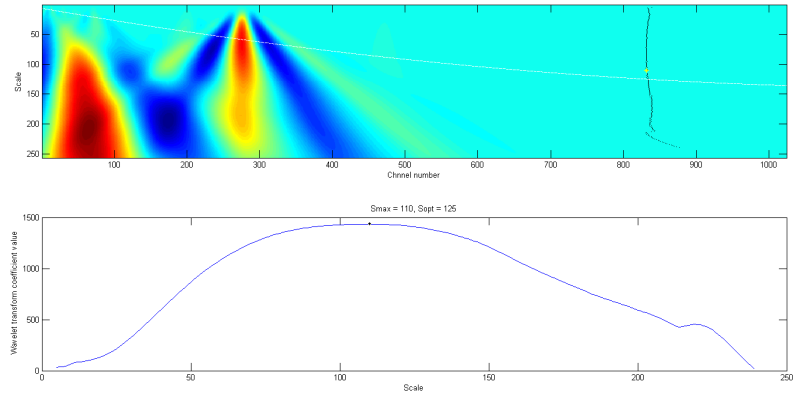


Figure 4.15: The missing Tl-208 peak in the Mn-54 spectrum.

Isotope : Co-60

Figure 4.16 shows the WTMM lines that remain in the collection after the vertical filter and before scale filtering with the DRF curve. The values of s_{max} along the WTMM lines were too much smaller than optimal scales for both 1173 keV peak and 1333 keV peak. Therefore both of the lines were filtered out and both characteristic peaks are missed for this isotope. This shows that further investigation of the DRF function needs to be carried out.

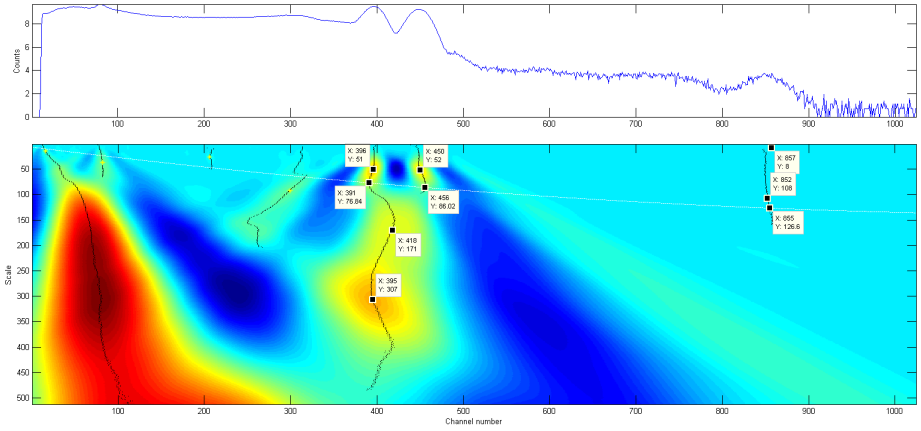


Figure 4.16: Features in the Co-60 spectrum that were identified prior to the scale filter.

Isotope : Na-22

The characteristic peak of 1275 keV was detected for Na-22 at the correct channel of 418, assuming true value given by Origin. The K-40 peak and the Tl-208 peak were missed because of the same reason stated for the missing Tl-208 peak in Am-241 spectrum (Figure 4.18 and Figure 4.19) The peak area was 11429592 by the algorithm deviating by 5.801% to Origin result 10802900.

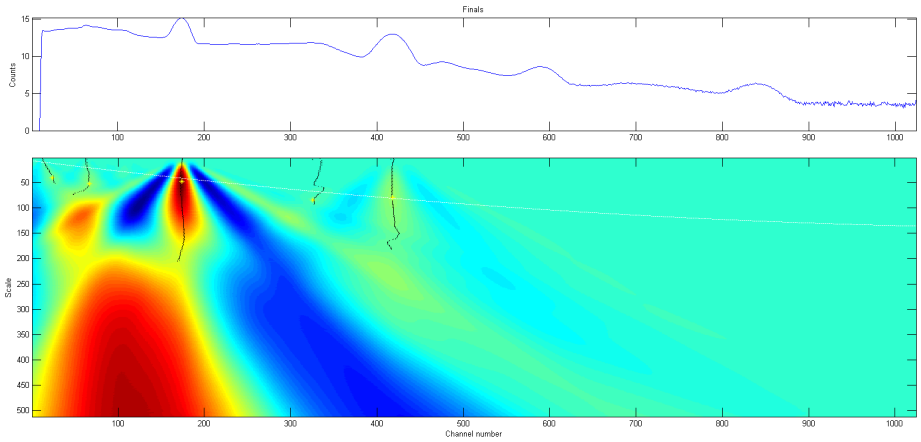


Figure 4.17: Features in the Na-22 spectrum that were identified.

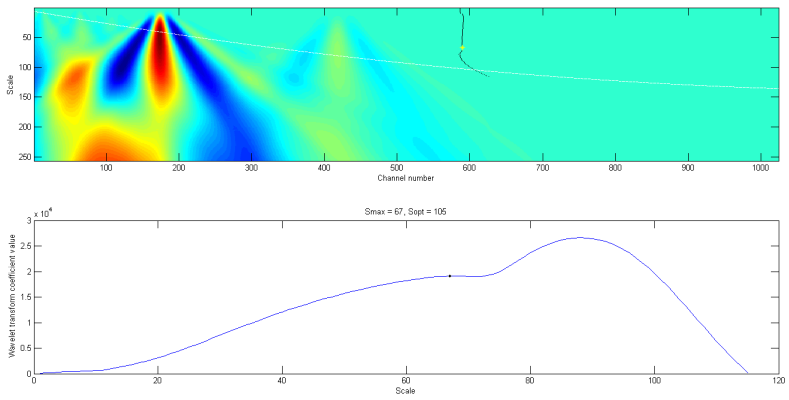


Figure 4.18: The missing K-40 peak in the Na-22 spectrum.

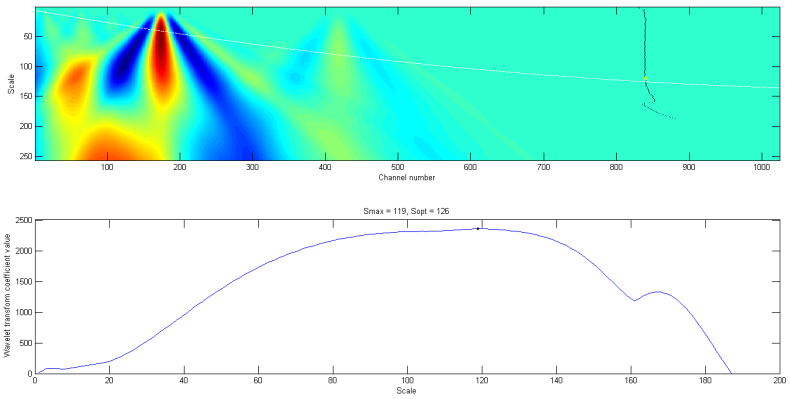


Figure 4.19: The missing K-40 peak in the Na-22 spectrum.

Up to this point, the algorithm has been evaluated for peak identification and quantification on the real spectra of selected isotopes. Except for the two in Co-60 (1.17 MeV And 1.33 MeV), all the characteristic photopeaks in the spectra are detected and quantified by the algorithm, which are the 59.5 keV in Am-241, the 122keV in Co-57, the 662 keV in Cs-137, the 835 keV in Mn-54, and the 1275 keV in Na-22. Occasionally photopeaks by background radiation such as the 1460 keV by K-40 and the 2614 keV by Tl-208 are detected as well. The peaks undetected are due to the design of filters related with scale and DRF. This should be addressed to in future study.

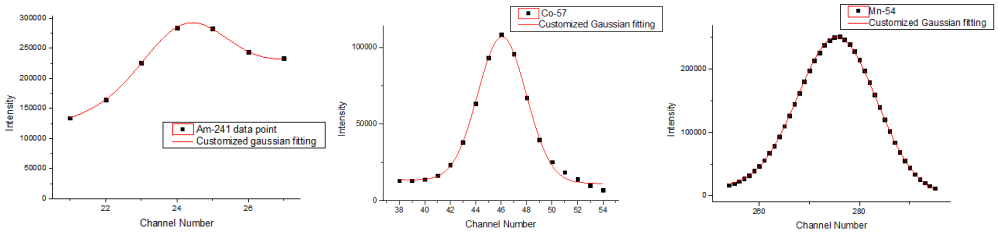
Table 4.4: Peak information obtained by wavelet analysis and NNLS calculation

	Centroid(ch)	Area
Am-141	24	1466709
Co-57	46	642414
Cs-137	223	508599
Mn-54	275	4771418
Co-60	N/A	N/A
Na-22	418	11429592

4.2.2 The results from Origin

In this section, data were fitted with Gaussian functions and a linear baseline in Origin to acquire peak centroids and peak areas as shown in Figure 4.20. Complete results are listed in Table 4.5.

.



(a) Peak of 59.5 keV in Am-241 spectrum

(b) Peak of 122 keV in Co-57 spectrum

(c) Peak of 835 keV in Mn-54 spectrum

Figure 4.20: Eamples of peak fitting results in Origin.

Table 4.5: Peak information obtained by Origin 9

Isotope	E(keV)	Peak centroid(ch)	FWHM(ch)	Peak Area	Standard Error for Peak Area
Am-241	59.5	24	2.73891	315787	8764
Co-57	122	46	4.47093	449963	10311
Cs-137	662	223	16.54186	542473	3258
Mn-54	835	276	18.49564	4830570	28650
Co-60 [1]	1173	397	22.4203	283840	2422
Na-22	1275	418	23.82546	10802900	41625
Co-60 [2]	1333	449	25.15259	274252	4020
K-40	1460	472	26.32335	74389	998
Tl-208	2614	834	38.39829	10407	400

4.2.3 Comparisons for subsection 4.2.1 and subsection 4.2.2

Comparison is made based on results from ORIGIN 9, as shown in Table 4.6. The background peaks for K-40 and Tl-208 were not detected in the background spectrum, therefore there is no information on those peaks. Using

Table 4.6: Comparison between solutions from proposed algorithm and Origin 9

Isotope	E(keV)	NNLS Centroid (ch)	NNLS Area	ORIGIN Centroid(ch)	ORIGIN Area	Deviation between Centroid	Deviation between Area
Am-141	59.5	24	1466709	24	315787	0.000%	364.461%
Co-57	122	46	642414	46	449963	0.000%	42.770%
Cs-137	662	223	508599	223	542473	0.000%	6.244%
Mn-54	835	275	4771418	276	4830570	0.362%	1.225%
Co-60	1173	N/A	N/A	397	283840	N/A	N/A
	1333	N/A	N/A	449	274252	N/A	N/A
Na-22	1275	418	11429592	418	10802900	0.000%	5.801%

the results from Origin as true values, except the two missing peaks of the Co-60 spectrum, all other characteristic peaks were detected with centroid deviation of less than 1%. However, the peak areas calculated by NNLS were not in compliance with those from Origin. Deviation of peak areas decreases significantly with energy. The reason of this may lie in the fact that low energy peaks sit on larger background and accumulated Compton continuum, making it difficult for Origin to accurately determine peak area. Peak areas calculated by NNLS were processed by wavelet extraction, therefore would suffer less influence from baseline and continuum.

As shown in the comparison, for single-peaks sitting on baseline or continuum that can be detected in the spectra by the wavelet-based algorithm, te algorithm provides accurate solutions for peak centroids and areas as good as Origin. However, better than Origin, with calibrated DRF of the detector, our algorithm can not only automatically resolve overlapping peaks but also can still provide overlapping peaks' centroids and areas.

Below is a part of Ba-133 spectrum, sitting on a simulated linear baseline. Origin can not resolve these four peaks at 276, 301, 355 and 383 keV. However

with feature extraction ability of wavelet transform, our algorithm still provides indications on peak centroids and peak areas, though with uncertainty to certain degree, which should be explored in the next step to accurately determine.

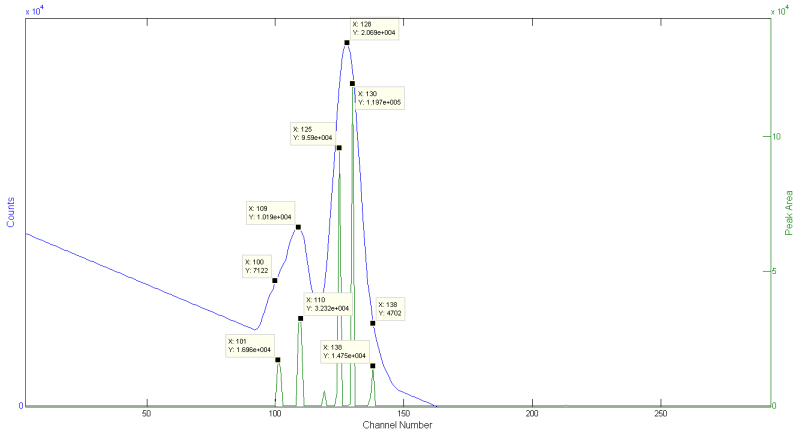


Figure 4.21: The peak areas determined by wavelet algorithm in the Ba-133 spectrum.

CHAPTER 5

EXPLORATION ON THE UNCERTAINTY OF PEAK AREA

5.1 Conventional Uncertainty of Peak Area

Conventionally, the uncertainty of a peak's area is calculated according to a region of interest, where boundaries of the region are appearing as abrupt rise and drop in counts number. The left bound and the the right bound of the peak region can be adjusted case by case. Mathematically, the uncertainty of peak area is determined by following equation:

$$\sigma_A = \sqrt{A + D(1 + \frac{n}{2m})}; \tag{5.1}$$

where A is net peak area (counts accumulated above an average background line below the net peak region), U is the upper boundary of the net peak region, L is the lower boundary of the net peak region, n is the number of channels within the net peak region, m is the number of channels of extension beyond the net peak region, D is the extended background area (counts accumulated below extension beyond the net peak region below background line) including background count integration from $L - m$ to $L - 1$ and from $U + 1$ to $U + m$, as shown in Figure 5.1 [1].

Our code implemented the equation above with extension width m = 5 channels, and the uncertainties were found for detectable peaks in real spectra, using the technique described in the following sections.

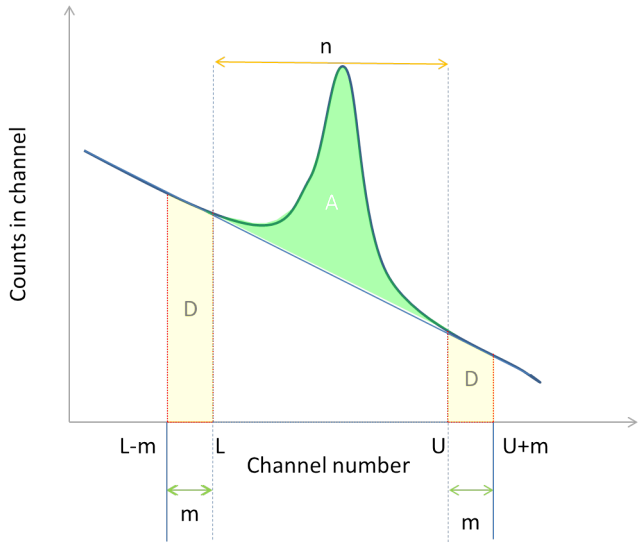


Figure 5.1: Definition of uncertainty of peak area in conventional way [1]

5.2 Attempt on Finding NNLS Uncertainty of Peak Area

The derivation of the uncertainty of a peak's area comes from error propagation in least square problem [31] [32] [33] [34] [29]. Since we are, in essence, evaluating the error in variables (the error of k in $S = Bk$), the answer has close relation with property of matrix B

$$S = Bk, \tag{5.2}$$

where B is the basis function matrix. The NNLS solution for vector k would be

$$k = (B^t B)^{-1} B^t S \tag{5.3}$$

$$= OS. \tag{5.4}$$

where

$$O = (B^t B)^{-1} B^t \tag{5.5}$$

and the minimum of $\|Bk - S\|$ is achieved by this value of k (please refer to Chapter 2).

Then covariance matrix for k is given by

$$\begin{aligned} C_k &= E\{[k - E[k]][k - E[k]]^t\} && (5.6)\\ &= E\{O[S - E[S]][S - E[S]]^t O^t\} && (5.7)\\ &= OE\{[S - E[S]][S - E[S]]^t\}O^t && (5.8)\\ &= OC_S O^t, && (5.9) \end{aligned}$$

where E is the expectation operator, and C_S is the covariance matrix of S [35] [36] [37]. In practical situations, C_S is not calculable because the basis matrix B is singular [38]. However by assuming the psuedo-inverse of the basis matrix B to be the inverse, we will have the generalized expression of uncertainty of k. Then we will deal with the special case of the singularity of the basis function from this generalized expression. So

$$C_S = \sigma_S^2 I. \tag{5.10}$$

Equation (5.9) becomes

$$\begin{aligned} C_k &= OC_S O^t && (5.11)\\ &= (B^t B)^{-1} B^t \sigma_S^2 I[(B^t B)^{-1} B^t]^t && (5.12)\\ &= \sigma_S^2 (B^t B)^{-1} B^t I[(B^t B)^{-1} B^t]^t. && (5.13) \end{aligned}$$

Since $B^t B$ is symmetrical, $(B^t B)^{-1}$ is also symmetrical too,

$$[(B^t B)^{-1}]^t = (B^t B)^{-1}. \tag{5.14}$$

Therefore

$$C_k = \sigma_S^2 (B^t B)^{-1} B^t B (B^t B)^{-1} \quad (5.15)$$

$$= \sigma_k^2 I = \sigma_S^2 (B^t B)^{-1}. \quad (5.16)$$

However, because of the nature of our basis function, it is significantly singular. For example, the specially designed basis function matrix B for a spectrum measured by the Ortec 2×2 NaI detector which has 1024 channels, has a rank of only 442, making it singular. However, the errors in k would only come from dimensions that are not free in the basis function matrix B. Therefore, the individual uncertainty of elements in the solution k would need to be scaled down to $\frac{1}{m-n}$, which is $\frac{1}{(1024-442)}$ in this case. With current references, we use the pseudo-inverse of $B^t B$ to calculate $(B^t B)^{-1}$. This step is achieved with the built-in function from Matlab: pinv(), which is the Moore-Penrose pseudo-inverse and is based on the theory of Single Value Decomposition(SVD).

5.3 Results of Both Methods from Section 5.1 and Section 5.2

Uncertainties were determined by both methods described in Section 5.1 and Section 5.2, and results are listed in Table 5.1. As listed in the table, uncertainties calculated from NNLS covariance matrix, conventional method and Origin significantly differ between each other, like they are living in parallel universe. A meaningful criteria or standard for uncertainty calculation needs to be established for further exploration of this problem.

Table 5.1: Comparison of uncertainty obtained from proposed algorithms

Isotope	NNLS Uncertainty	Conventional Uncertainty	ORIGIN Uncertainty	Deviation between Uncertainty (NNLS and ORIGIN)	Deviation between Uncertainty (NNLS and conventional)	Deviation between Uncertainty (conventional and Origin)
Am-141	613	1665	8764	93.005%	63.177%	81.005%
Co-57	1368	850	10311	86.733%	60.968%	91.758%
Cs-137	788	761	3258	75.813%	3.585%	76.651%
Mn-54	767	2282	28650	97.323%	66.391%	92.034%
Co-60	N/A	N/A	2422	N/A	N/A	N/A
	N/A	N/A	4020	N/A	N/A	N/A
Na-22	692	3409	41625	98.338%	79.701%	91.810%

CHAPTER 6

CONCLUSIONS AND FUTURE WORK

6.1 Conclusion

This thesis focuses on answering the problem of peak detection and peak quantification of the peak centroid, area and the uncertainty of the area. The continuous wavelet transform and NNLS were utilized in designing the algorithm. Comparing with reference spectra or results from other software tools, the performance of the algorithm was evaluated on both simulated signals and real spectra.

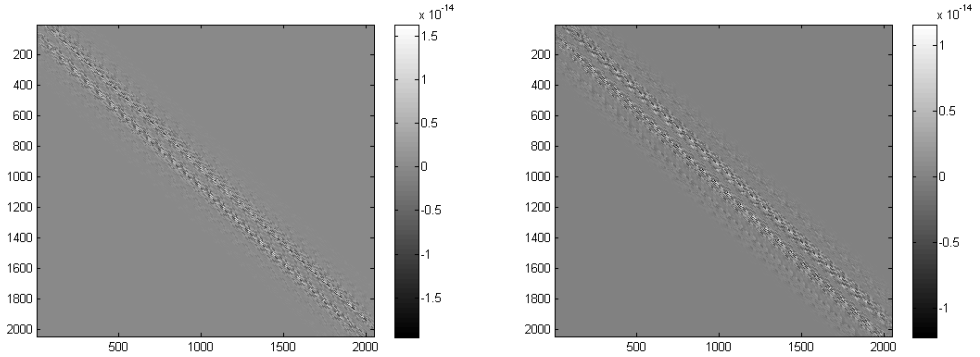
On the assessment of results for real spectra, the algorithm works for finding the characteristic peaks in Am-241, Co-57, Cs-137, Mn-54 and Na-22 but misses the two in Co-60. Preliminary interpretation on this problem is that the Compton continuum of 1.33 MeV peak raised the baseline on which 1.17 MeV peak sits on. Therefore FWHMs from those two peaks used for DRF calibration is broader than it should be. It is reasonable to rethink the calibration data of DRF curve on these two peaks. If the FWHM gets smaller, the s_{opt} from the fitted DRF could be smaller, making the detection of Tl-208 peak a possible task as well, because the difference between s_{max} and s_{opt} gets smaller. For peaks identifiable by the algorithm, centroids and areas are assessed comparing with results get from Origin. Uncertainties of peak areas are calculated in both conventional ways and using NNLS solution analysis. Comparisons were made between the NNLS solutions and Origin fitting errors. Some deviations were large but there are directions for future works that can be concluded from these results.

6.2 Future work

There are many aspects that need to improve in order to better the performance of current algorithm. The first one is the variability of basis function matrices. Being one of the two most important elements in performing NNLS computation, the basis matrix B must be carefully determined. Attempts were made to compute this matrix on different computers to calculate the results as quick as possible. It was discovered that sometimes the computations of B from different computers do not match. It does not mean that every element in the corresponding position of different matrices varies with each significantly. For example, the basis function matrices generated for simulated peak of 65.355 channels in FWHM with wavelet 'bior2.6' from two computers are shown in Figure. They look the same and the results calculated for peak areas are the same, but 3,647,479 elements out of total 2048 $\times$ 2048 are different. For wavelet 'gaus4', the situation gets worse: even the results calculated for peak areas are different with the basis function matrices calculated from different computers. Since basis function matrices are generated with the same wavelet at the same scale, they should not vary according to different computers. This might need more inspection considering the inaccuracy of computation would be compounded as a result.

The second aspect is on what scale to solve the NNLS problem. Sometime because the position of WTMM lines varies greatly along scales, uncertainty in peak centroids can be large. For these situations, the proper way of determining which scale should be used to solve the NNLS problem needs to be found.

The variability in WTMM line positions also brings up the third need for improvement: further filters for true photopeak determination. Compton edges are sometimes detected as peaks in the wavelet extraction part. However the WTMM lines associated with Compton edges can have more variabil-



(a) Difference matrix of Basis matrices for 'bior2.6' wavelet

(b) Difference matric of Basis matrices for 'gaus4' wavelet

Figure 6.1: Difference matrix of basis function matrices calculated from different PCs

ity in scale positions than true peaks. Distinguishing features like Compton edges can be accomplished by designing new filters for them. Moreover, only WTMM lines in terms of scales are created as filters for uptodate algorithm, attempts can be made to create WTMM lines with respect to energy channels. This extra filter can offer additional reference for peak searching.

Another aspect of future work is handling boundary conditions of the carved-out small piece of the spectrum. As was shown in Chapter 4, a preliminary way of handling boundary effect was used to determine peak areas. In simple, simulated signal analysis this technique works, but it is not sufficient in more complex situations. Advanced techniques should be tested to evaluate the efficacy of this approach.

There is also a need to simulate the situation where peaks are sitting on non-linear baseline and evaluate peak area calculation in that situation.

Sixth, continuous exploration needs to be carried out on quantification on uncertainty of peak area. There are commercial software built-in with products and packages that possess the ability to quantify peak area uncertainty. But the current standard of quantification on uncertainty of peak area needs to be found to evaluate the performance of our algorithm. Further research should be carried out here.

Next prospect of improvement is in careful data collection and curve fitting for DRF function. The DRF for NaI scintillators has an empirical non-linear curve template that can be fit [3] [39] [40]. However, data to fit this DRF function can be difficult to control in a consistently-reproducible environment. Future data collection needs to be carefully monitored to maintain the calibration environment so that we can have least interference of other variables in the production of DRF function. The template of the DRF for NaI needs continuous updating.

Also, as mentioned in Chapter 2, several near symmetric wavelets were used in the beginning of study. Both fourth order of Gaussian function and Coiflet of order two wavelets were not giving good enough results when solving for peak areas of simulated signals. Therefore results and figures presented here are generated by bi-orthogonal wavelet of order 2.6 solely. Yet even the bi-orthogonal wavelet might not be perfect with respect to minimizing the closest resolvable distance. For instance, as discussed in Chapter 4, the closest resolvable distance between two equal area peaks was found to be 16 channels. It is possible that new custom-designed wavelets may be better suited to solve this problem [4].

Last but not the least, the software written for this research has quite a few parameters that need to be optimized, either for the purpose of efficiency or the purpose of accuracy. Such parameters are the threshold value for finding the local maxima in scalogram, the lower and upper bound of the carved-out spectrum interval for NNLS calculation, the tolerance of the difference between maximum scale and the optimal scale of the WTMM line.

APPENDIX A

MATLAB FUNCTIONS

```
function [arena,rec]=peaks2(spect0,wavelet,scales)
```

==Data Input and pre-processing

```
load('final.mat');
spect0=spect0';
numch=size(spect0);chl=max(numch);
ch=1:chl;
```

==Direct Localmaxima

```
wt=cwt(spect0,scales,wavelet);
y=jmax(wt);[irow,icol]=find(y);
figure();
subplot('position',[0.05 0.7 0.92 0.25]); plot(spect0);axis tight;
hold on;subplot('position',[0.05 0.04 0.92 0.6]);
imagesc(wt);
hold on
plot(ch(icol),scales(irow),'k.','markersize',4);
plot(SC,'w.','markersize',4);title('JMAX');
set(0,'DefaultFigureWindowStyle','docked')

y2=filterH(wt,y);
[irow2,icol2]=find(y2);
figure();
subplot('position',[0.05 0.7 0.92 0.25]);plot(log(spect0+1));axis tight;
hold on;subplot('position',[0.05 0.04 0.92 0.6]);
imagesc(wt); hold on;plot(ch(icol2),scales(irow2),'k.','markersize',4);
plot(SC,'w.','markersize',4);title('filterH')
```

```
set(0,'DefaultFigureWindowStyle','docked')

z=lines4(y2);
z2=linesplit(z);
```

==DRF check

```
for i=1:max(z2(:,1))
    dummyIndex=find(z2(:,1)==i);
    z2dummy=z2(dummyIndex,:);
    iden=polyxpoly(ch,SC,z2dummy(:,3),z2dummy(:,2))
    if isempty(iden)
        z2(dummyIndex,1)=0;
    end
end
z3=z2((z2(:,1)~=0),:);
```

==Get Optimal Scales and Rearrange Line Numbers

```
z5=flipud(z3);
d=unique(z5(:,1));len=length(d);
ops=zeros(len,1);opt=zeros(len,1);finalseed=[];seedid=[];

for i=1:len
    z5dummy=z5(find(z5(:,1)==d(i)),:);
    opt(i)=round(max(polyxpoly(ch,SC,z5dummy(:,3),z5dummy(:,2))));
    ops(i)=SC(opt(i));
end

Ltwist=[opt d];
dumid=(1:max(size(Ltwist)))';Ltwist=[Ltwist dumid];Ltwist=sortrows(Ltwist);
Ltwist=[ Ltwist(:,2) Ltwist(:,3) Ltwist(:,1)];
```

==Start Line by Line Inspection [Eliminations by Scales and Slopes]

```
maxs=zeros(len,1);cen=zeros(len,1);rnorm=zeros(len,1);k5=zeros(len,chl);
```

```
area=zeros(len,1);uncer=zeros(len,1);rec=[];
for i=1:len
dd=z5(find(z5(:,1)==d(i)),:);
maxs(i)=maxtesting5(dd,wt,i);

 if maxs(i)==0
     continue
 else
      cen(i)=dd(find(dd(:,2)==maxs(i)),3);
 end
   if spect0(cen(i))==0
      continue
   end
  if maxs(i)>=(ops(i)-6)
    low=floor(cen(i)-3*floor(fh(cen(i))));
    up=ceil(cen(i)+3*ceil(fh(cen(i))));
    lowa=floor(cen(i)-0.6*floor(fh(cen(i))));
    upa=ceil(cen(i)+0.6*ceil(fh(cen(i))));
    if low>0&&up<chl
        n5=spect0(low:up);
        nnlsk5=[zeros(1,low+1001),n5,zeros(1,chl-1-up+1000)];
    elseif low<=0
        low=1;
        lefth=cen(i);
        up=2*lefth-1;
        n5=spect0(low:up);
        if length(n5)<=10
            continue
        end
        nnlsk5=[zeros(1,low+1001),n5,zeros(1,chl-1-up+1000)];
        if lowa<=0
            lowa=1;upa=2*cen(i)-1;
        end
    elseif up>=chl
        up=chl;
        righth0=cen(i);
        righth=length(n5)-righth0;
        low=chl-2*righth+1;
```

```
        n5=spect0(low:up);
        if length(n5)<=10
            continue
        end
        nnlsk5=[zeros(1,low+1001),n5,zeros(1,chl-1-up+1000)];
        if upa>=chl
            upa=chl;lowa=chl-2*cen(i)+1;
        end
    end
        [k5(i,:),rnorm(i)]=
        quickNNLSchl(n5,nnlsk5,'bior2.6',Bb1024,
        ops(i),maxs(i),cen(i),chl,low,up);
        area(i)=sum(k5(i,lowa:upa));uncer(i)=0;
        if isempty(finalseed)
            finalseed=d(i);seedid=i;
        else
            finalseed = cat(2,finalseed,d(i));seedid=cat(2,seedid,i);
        end
   end
 end
```

==Final Output and Display

```
arena0=[cen'; area' ;uncer']';
arena=arena0((arena0(:,2)~=0),:);
z6=z5;
XOR=setxor(finalseed,z6(:,1));
lex=length(XOR);
for i=1:lex
    fifi=find(z6(:,1)==XOR(i));
    z6(fifi,1)=0;
end
z6=z6(z6(:,1)~=0,:);
LIN=Ltwist;
fifi4=[];
for i=1:length(seedid)
    eee=LIN(LIN(:,2)==seedid(i),:);
    if isempty(fifi4)
```

```
    fifi4=eee;
    else fifi4=[fifi4;eee];
    end
end
fifi4=[fifi4(:,3) fifi4(:,1) fifi4(:,2)];
fifi4=sortrows(fifi4);
fifi4=fifi4(:,3);
figure(251);
subplot('position',[0.05 0.7 0.92 0.25]);
plot(log(1+spect0));
hold on;
axis tight;
title('Finals');
hold off;
hold on;
subplot('position',[0.05 0.04 0.92 0.6]);
imagesc(wt);
hold on;
if isempty(seedid)
    hold off
else
plot(ch(z6(:,3)),scales(z6(:,2)),
'k.','markersize',4);
plot(ch(cen(seedid)),scales(maxs(seedid)),
'y*','markersize',6);
plot(SC,'w.','markersize',4);
title(num2str(fifi4'));
hold off;
end
set(0,'DefaultFigureWindowStyle','docked')
            cen0=cen(cen(:)~=0);
            maxs0=maxs(maxs(:)~=0);
            LIN=Ltwist;
            LINSEED=find(cen(:)~=0);
            StarInd=d(LINSEED);

            z7=z5;
            XOR2=setxor(StarInd,z7(:,1));
```

```
            lex2=length(XOR2);
            for i=1:lex2
            fifi2=find(z7(:,1)==XOR2(i));
            z7(fifi2,1)=0;
            end

            z7=z7(z7(:,1)~=0,:);

fifi3=[];
for i=1:length(StarInd)
    qq=LIN(LIN(:,1)==StarInd(i),:);
    if isempty(fifi3)
    fifi3=qq;
    else fifi3=[fifi3;qq];
    end

end
    fifi3=[fifi3(:,3) fifi3(:,1) fifi3(:,2)];
    fifi3=sortrows(fifi3);
    fifi3=fifi3(:,3);
            figure(250);
            subplot('position',[0.05 0.7 0.92 0.25]);
            plot(log(1+spect0));
            hold on;
            axis tight;
            title('After Straightness Filter');
            hold off;
            hold on;
            subplot('position',[0.05 0.04 0.92 0.6]);
            imagesc(wt);
            hold on;
            plot(ch(z7(:,3)),scales(z7(:,2)),'k.','markersize',4);
            plot(ch(cen0),scales(maxs0),
            'y*','markersize',6);
            plot(SC,'w.','markersize',4);
            title(num2str(fifi3'));
            hold off;
            set(0,'DefaultFigureWindowStyle',
```

```
    'docked')
```

==User Interaction

```
while(1)
    diskey0=
    input(['----------------------
    Choose line number from 1 to ',
    num2str(len),
    ' to display or choose 0 to exit : '],
    's');
    diskey=str2double(diskey0);
    if diskey>=1 && diskey<=len
        dd=z5(find(z5(:,1)==d(diskey)),:);
        figure();
        subplot(212)
        t=dd(:,2);r=wt(dd(:,4));
        if maxs(diskey)~=0
            ds=find(t==maxs(diskey));
            fun=dd(find(dd(:,2)==maxs(diskey)),3);
            plot(t,r,'b',maxs(diskey),r(ds),'k.');
            title([num2str(diskey),
            ' Smax = ',
            num2str(maxs(diskey)),
            ', Sopt = '
            ,num2str(ops(diskey)),
            ', centroid around ',
            num2str(cen(diskey))]);
            subplot(211);
            imagesc(wt);
            hold on;
            plot(ch(dd(:,3)),
            scales(dd(:,2)),
            'k.','markersize',4);
            plot(ch(fun),
            scales(maxs(diskey)),'y*',
            'markersize',6);
            plot(SC,'w.','markersize',4);
```

```
            title(['line',num2str(diskey)]);

        else
            plot(t,r,'b');
             title([num2str(diskey),
             ' Smax = ',
             num2str(maxs(diskey)),
             ', Sopt = ',
             num2str(ops(diskey)),
             ', centroid around ',
             num2str(cen(diskey))]);
            subplot(211);
            imagesc(wt);
            hold on;
            plot(ch(dd(:,3)),
            scales(dd(:,2)),'k.',
            'markersize',4);
            plot(SC,'w.','markersize',4);
            title(['line',num2str(diskey)]);
        end
        set(0,'DefaultFigureWindowStyle',
        'docked')

    elseif diskey==0
        break
    else
    end
end

arena
```

```
function [y,I] = jmax(x)

r = size(x,1);
y = [zeros(r,1) diff(abs(x),1,2)];
y(abs(y)<sqrt(eps)) = 0;
y(y<0) = -1;
y(y>0) = 1;
y = diff(y,1,2);
I = find(y==-2);
y = zeros(size(x));
y(I) = 1;

m=max(max(x));
I=find(y);
threshold=-10; %2.8 for log space of other wavelets %~50 normal space
for i=1:size(I,1)
    if x(I(i))<threshold
        y(I(i))=0;
    end
end
```

```
function [ y ] = filterH( x,y )

%FILTERH Summary of this function goes here
%   Detailed explanation goes here

r2 = size(x,1);
r1 = size(x,2);
s=size(x);
w=30; %width to filter over
I=find(y);
y2=zeros(size(y));

for i=1:size(I,1)
    [a,b]=ind2sub(s,I(i));
    if (b-w)>0
        if (b+w)<r1
            m=max(x(a,(b-w):(b+w)));
            if x(a,b)< m
                y(a,b)=0;

                for j=(b-w):(b+w)
                   if abs(m-x(a,j))<eps
                       y2(a,j)=1;
                   end
                end

            end
        else
            m=max(x(a,(b-w):r1));
            if x(a,b)< m
                y(a,b)=0;

                for j=(b-w):(r1)
                   if abs(m-x(a,j))<eps
                       y2(a,j)=1;
                   end
                end

            end
```

```
        end
    else
        m=max(x(a,1:(b+w)));
        if x(a,b)< m
                y(a,b)=0;

                for j=1:(b+w)
                   if abs(m-x(a,j))<eps
                       y2(a,j)=1;
                   end
                end

            end


    end
end


%y=y+y2;


end
```

```
function [ y ] = lines4( y )

% LINES4
% Given a matrix y of zeros and ones, changes the values of ones to the
% position nearest maxima above it. End of maxima chains are set to a value
% of -1 instead.
%
% Algorithm:
% 1. Find all ones on the bottom row of the matrix.
% 2. For each one, find the "nearest neighbor above".
%       That is, look up one row. Is there another one within a set
%       horizontal difference (hw)? If so, set the value of the lower one
%       to the index value of the upper one.  If not, look up another row.
%       Repeat if needed up to a maximum vertical difference of hw.  If no
%       one is found, set the value of the lower one to '-1' to signify the
%       end of a chain.
% 3. Repeat for each row.
% 4. At the end, there are still ones in the output matrix. These
% correspond either to the top most point of the line (the end of the
% chain) or isolated maxima points.  We set all of these values to '-1'.
%
%
% This is the final version of this code.


%Note that when referring to the elements of this matrices by the normal
%[a,b] index, the first index corresponds to the y-position (the row
%number) and the second corresponds to the x-position.  Then if we know
%what (x,y) coordinate we want, we actually need to call it as [y,x].  This
%is why some of these variables are named 'sx' or 'sy': to remind that they
%are the size in that direction, etc
sx=size(y,2); %size in x direction
sy=size(y,1); %size in y direction
s=size(y);    %size

%how spaced should we allow maxima to be?
hwl=10; %horizontal radius to check.
vwl=5;  %vertical radius to check
```

```
%scan across each row, starting from the bottom
%Note: loop ends at 2nd row. It is not possible to look for maxima above
%the first row, so it is handled separately.
for i=sy:-1:2
    Il=find(y(i,:)); %(lower row)
    yl=i; %
    if numel(Il)~=0  %make sure there is at least one maxima
    %in a row before checking
        for ii=1:size(Il,2)
            xl=ind2sub(s,Il(ii));

            vl=min(i-1,vwl); %vertical limit to check
            hlLEFT=min([xl-1,hwl]);
            %horizontal limit, must stay within bounds
            hlRIGHT=min([hwl,sx-xl]);
            for H=1:vl
            %check next row first,
            %then so on if none found

                %set value to -1
                %if there is no maxima above it,
                %i.e. it's the
                %end of the chain
                temp=y(i-H,(xl-hlLEFT):(xl+hlRIGHT));
                %ind=find(temp);
                ind1=find(temp==1);
                if numel(ind1)==0 %no ones found
                    if H==vl %no rows to check
                        y(yl,xl)=-1; %-1 signifies end of chain
                        break
                    end
                else %new value found value found

                    [dy,dx]=ind2sub(size(temp),ind1);
                    y(yl,xl)=sub2ind(s,yl-H,xl-hlLEFT+dx-1);
                    break;

                end
```

```
            end
        end
    end
end

%Fix all of the ones on the topmost row (they are the ends of maxima
%chains or isolated maxima)
y(find(y==1))=-1;
end
```

```
function [ v ] = linesplit( y )

%LINESPLIT Takes the output of lines4 and separates the lines.
% returns 4 columns: line number, two normal subscripts, and the index.

sx=size(y,2); %size in x direction
sy=size(y,1); %size in y direction
s=size(y);    %size

n=0;
v1=[]; %line number
v2=[]; %first index
v3=[]; %second index
v4=[];
for i=sy:-1:2
    Il=find(y(i,:)); %(lower row)
    yl=i; %
    if numel(Il)~=0
    %make sure there is at least one maxima
    %in a row before checking
        for ii=1:size(Il,2)
        %find returns a 1xn vector
            n=n+1; %next line

            iii=sub2ind(s,i,Il(ii));
            while iii>0
                [a,b]=ind2sub(s,iii);
                v1=[v1 n];

                v2=[v2 a];
                v3=[v3 b];
                v4=[v4 iii];
                iii=y(a,b); %next value
                y(a,b)=0; %get rid of value so it isn't checked again
            end
        end
    end

v=[v1' v2' v3' v4'];
```

```
end
```

```
function [maxind1,maxind]=maxtesting5(t,wt,lineind)

load('final.mat');
t0=t(:,2);
r=wt(t(:,4));



x=r;
n=max(size(x));
maxlist=[];ds=[];
d=5;
for i=2:(n-1)
    m=[];
    if i<=d
        if (i+d)>=length(x)
            m=max(x(1:end));
        else
            m=max(x((1):(i+d)));
        end
        if abs(x(i)-m)<eps
            maxlist=[maxlist t0(i)];
            ds=[ds i];
        end
    else
        if i>=(n-d)
            % m=max(x((i-d):(i)))
            m=max(x((i-d):(n)));
            if abs(x(i)-m)<eps
                maxlist=[maxlist t0(i)];
                ds=[ds i];
            end
        else
            m=max(x((i-d):(i+d)));
            if abs(x(i)-m)<eps
                maxlist=[maxlist t0(i)];
                ds=[ds i];
            end
        end
    end
```

```
end
%maxlist
% maxind1=maxind(1);
if isempty(maxlist)
    maxlist=0;
end
maxind1=maxlist(1);
```

```
function [k,resnorm]=
quickNNLSchl(n5,y,wavelet,B,ops,maxs,cen,chl,low,up)


ch=1:chl;

c1=cwt(y,1:512,wavelet);
cc1=c1(maxs,(1002:(1002+low+length(n5)+chl-up-2)));
[k,resnorm]=lsqnonneg(B,cc1');
 y0=y(1002:(1002+low+length(n5)+chl-up-2));
 [AX,H1,H2] = plotyy(ch,y0,ch,k');
title([' Smax = ',num2str(maxs),
', Sopt = ',num2str(ops),
', centroid around ',num2str(cen)])
set(get(AX(1),'Ylabel'),'String','Counts')
set(get(AX(2),'Ylabel'),'String','Peak Area')

xlabel('Channel Number','FontSize',12)

set(get(AX(1),'Ylabel'),'FontSize',12);
set(get(AX(2),'Ylabel'),'FontSize',12);


set(AX(1), 'xlim', [1 chl])
set(AX(2), 'xlim', [1 chl])


end
```

REFERENCES

[1] G. R. Gilmore, *Practical gamma-ray spectrometry, 2nd Edition.* Chichester, England: Wiley, 2008.

[2] T. Burr and M. Hamada, "Radio-isotope identification algorithms for nai spectra," *Algorithms*, vol. 2, no. 1, pp. 339–360, 2009.

[3] G. F. Knoll, *Radiation Detection and Measurement.* New York: Wiley, 1979.

[4] M. A. Mariscotti, "A method for automatic identification of peaks in the presence of background and its application to spectrum analysis," *Nuclear Instruments and Methods*, vol. 50, no. 2, pp. 309–320, 1967.

[5] S. Ihantola, A. Pelikan, R. Pllnen, and H. Toivonen, "Advanced alpha spectrum analysis based on the fitting and covariance analysis of dependent variables," *Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 656, no. 1, pp. 55–60, 2011.

[6] P. Du, W. A. Kibbe, and S. M. Lin, "Improved peak detection in mass spectrum by incorporating continuous wavelet transform-based pattern matching," *Bioinformatics*, vol. 22, no. 17, pp. 2059–2065, 2006.

[7] G. Strang and T. Nguyen, *Wavelets and Filter Banks.* Wellesley, MA: Wellesley-Cambridge Press, 1997.

[8] S. G. Mallat, *A Wavelet Tour of Signal Processing, 2nd Edition.* San Diego: Academic Press, 1999.

[9] A. Caciolli, M. Baldoncini, G. P. Bezzon, C. Broggini, G. P. Buso, I. Callegari, T. Colonna, G. Fiorentini, E. Guastaldi, F. Mantovani, G. Massa, R. Menegazzo, L. Mou, C. R. Alvarez, M. Shyti, A. Zanon, and G. Xhixha, "A new fsa approach for in situ γ ray spectroscopy," *Science of the Total Environment*, vol. 414, pp. 639–645, 2012.

[10] C. Kuo, B. W. Reutter, and R. H. Huesman, "Resolution of the spectral technique in kinetic modeling," in *Proceedings of SPIE - The International Society for Optical Engineering*, vol. 4321, 2001, pp. 12–21.

[11] K. Kupfer, *Electromagnetic Aquametry Electromagnetic Wave Interaction with Water and Moist Substances.* Dordrecht: Springer: Springer, 2006.

[12] K. P. Whittall and A. L. MacKay, "Quantitative interpretation of nmr relaxation data," *Journal of Magnetic Resonance (1969)*, vol. 84, no. 1, pp. 134–152, 1989.

[13] G. Gambarota, B. Cairns, C. Berde, and R. Mulkern, "Osmotic effects on the $t-2$ relaxation decay of in vivo muscle," *Magnetic Resonance in Medicine*, vol. 46, no. 3, pp. 592–599, 2001.

[14] F. Gullo, G. Ponti, A. Tagarelli, G. Tradigo, and P. Veltri, "A time series approach for clustering mass spectrometry data," *Journal of Computational Science*, vol. 3, no. 5, pp. 344–355, 2012.

[15] V. Spyropoulou, M. A. Rapsomaniki, K. Theofilatos, S. Papadimitriou, S. Likothanassis, A. Tsakalidis, and S. Mavroudi, "Computational methods and algorithms for mass spectrometry based differential proteomics: Recent advances, perspectives and open problems," *Current Proteomics*, vol. 9, no. 3, pp. 143–159, 2012.

[16] B. Y. Renard, M. Kirchner, H. Steen, J. A. J. Steen, and F. A. Hamprecht, "Nitpick: Peak identification for mass spectrometry data," *BMC Bioinformatics*, vol. 9, 2008.

[17] C. Sullivan, M. E. Martinez, and S. Garner, "Wavelet analysis of sodium iodide spectra," *IEEE Transactions on Nuclear Science*, vol. 53, no. 5, pp. 2916–2922, 2006.

[18] C. J. Sullivan, S. E. Garner, K. B. Blagoev, and D. L. Weiss, "Generation of customized wavelets for the analysis of -ray spectra," *Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 579, no. 1, pp. 275–278, 2007.

[19] J. P. Li, in *The proceedings of the International Computer Congress 2004 on Wavelet Analysis and its Applications, and Active Media Technology*, Singapore, 2004.

[20] K. Lau and H. Weng, "Climate signal detection using wavelet transform: how to make a time series sing," *Bulletin - American Meteorological Society*, vol. 76, no. 12, pp. 2391–2402, 1995.

[21] S. Law and X. Zhu, *Damage models and algorithms for assessment of structures under operating conditions*. Boca Raton: CRC Press, 2009.

[22] A. Prokoph and F. Agterberg, "Wavelet analysis of well-logging data from oil source rock, egret member, offshore eastern canada," *AAPG Bulletin*, vol. 84, no. 10, pp. 1617–1632, OCT 2000, pT: J; NR: 35; TC: 14; J9: AAPG BULL; PG: 16; GA: 365MH; UT: WOS:000089954700006.

[23] K. He, C. Xie, S. Chen, and K. K. Lai, "Estimating var in crude oil market: A novel multi-scale non-linear ensemble approach incorporating wavelet analysis and neural network," *Neurocomputing*, vol. 72, no. 1618, pp. 3428 – 3438, 2009.

[24] K. R. Coombes, S. Tsavachidis, J. S. Morris, K. A. Baggerly, M. . Hung, and H. M. Kuerer, "Improved peak detection and quantification of mass spectrometry data acquired from surface-enhanced laser desorption and ionization by denoising spectra with the undecimated discrete wavelet transform," *Proteomics*, vol. 5, no. 16, pp. 4107–4117, 2005.

[25] R. Kronland-Martinet, "Wavelet transform for analysis, synthesis, and processing of speech and music sounds," *Computer Music Journal*, vol. 12, no. 4, pp. 11–20, 1988.

[26] I. Daubechies, "Ten lectures on wavelets." Philadelphia, Pa.: Society for Industrial and Applied Mathematics, 1992.

[27] C. L. Lawson and R. J. Hanson, *Solving least squares problems.* Englewood Cliffs, N.J.: Prentice-Hall, 1974.

[28] W. H. Press, *Numerical Recipes in C: The Art Of Scientific Computing.* Cambridge [u.a.]: Cambridge Univ. Press, 1990.

[29] H. Zhu, G. Leus, and G. B. Giannakis, "Sparsity-cognizant total least-squares for perturbed compressive sampling," *IEEE Transactions on Signal Processing*, vol. 59, no. 5, pp. 2002–2016, 2011.

[30] P. S. Addison, "Wavelet transforms and the ecg: A review," *Physiological Measurement*, vol. 26, no. 5, pp. R155–R199, 2005.

[31] A. E. Isbell, "Modeling error propagation in extinction coefficient calculations due to experimantal errors in the measured input data set," M.S. thesis, University of Illinois at Urbana-Champaign, Urbana-Champaign, 1990.

[32] J. Billeter, Y. . Neuhold, L. Simon, G. Puxty, and K. Hungerbhler, "Uncertainties and error propagation in kinetic hard-modelling of spectroscopic data," *Chemometrics and Intelligent Laboratory Systems*, vol. 93, no. 2, pp. 120–131, 2008.

[33] K. Faber and B. R. Kowalski, "Propagation of measurement errors for the validation of predictions obtained by principal component regression and partial least squares," *Journal of Chemometrics*, vol. 11, no. 3, pp. 181–238, 1997.

[34] P. R. Bevington and D. K. Robinson, *Data reduction and error analysis for the physical sciences.* Boston: McGraw-Hill, 2003.

[35] R. L. Plackett, *Principles of Regression Analysis.* Oxford: Clarendon Press, 1960.

[36] B. W. Rust and W. R. Burrus, *Mathematical programming and the numerical solution of linear equations.* New York: American Elsevier Pub. Co., 1972.

[37] H. Scheff, *The Analysis Of Variance.* New York: Wiley, 1959.

[38] P. A. Jansson, *Deconvolution of Images and Spectra, 2nd Edition.* San Diego: Academic Press, 1997.

[39] A. Cengiz, "An approximation for response function to -rays of nai(tl) detectors up to 1.5 mev," *Applied Radiation and Isotopes*, vol. 66, no. 10, pp. 1371–1376, 2008.

[40] L. Meng and D. Ramsden, "An inter-comparison of three spectral-deconvolution algorithms for gamma-ray spectroscopy," *IEEE Transactions on Nuclear Science*, vol. 47, no. 4 PART 1, pp. 1329–1336, 2000.

[41] S. M. Kay and S. L. Marple Jr., "Spectrum analysis - a modern perspective." *Proceedings of the IEEE*, vol. 69, no. 11, pp. 1380–1419, 1981.

[42] J. Billeter, Y. . Neuhold, and K. Hungerbhler, "Kinetic hard-modelling and spectral validation of rank-deficient spectroscopic data: A case study," *Chemometrics and Intelligent Laboratory Systems*, vol. 98, no. 2, pp. 213–226, 2009.

[43] J. Gill and G. King, "What to do when your hessian is not invertible: Alternatives to model respecification in nonlinear estimation," *Sociological Methods and Research*, vol. 33, no. 1, pp. 54–87, 2004.

[44] M. Slawski, R. Hussong, A. Tholey, T. Jakoby, B. Gregorius, A. Hildebrandt, and M. Hein, "Isotope pattern deconvolution for peptide mass spectrometry by non-negative least squares/least absolute deviation template matching," *BMC Bioinformatics*, vol. 13, no. 1, 2012.

[45] L. Li and T. P. Speed, "Parametric deconvolution of positive spike trains," *Annals of Statistics*, vol. 28, no. 5, pp. 1279–1301, 2000, cited By (since 1996):12. [Online]. Available: www.scopus.com

[46] D. B. Percival and A. T. Walden, *Wavelet Methods for Time Series Analysis*. Cambridge [u.a.]: Cambridge Univ. Press, 2000.

[47] J. F. Muzy, E. Bacry, and A. Arneodo, "Multifractal formalism for fractal signals: The structure-function approach versus the wavelet-transform modulus-maxima method," *Physical Review E*, vol. 47, no. 2, pp. 875–884, 1993.

[48] J. Stutz and U. Platt, "Numerical analysis and estimation of the statistical error of differential optical absorption spectroscopy measurements with least-squares methods," *Applied Optics*, vol. 35, no. 30, pp. 6041–6053, 1996.

[49] J. B. Ghasemi, Z. Heidari, and A. Jabbari, "Toward a continuous wavelet transform-based search method for feature selection for classification of spectroscopic data," *Chemometrics and Intelligent Laboratory Systems*, vol. 127, pp. 185–194, 2013.

[50] I. DAUBECHIES, "Orthonormal bases of compactly supported wavelets," *Communications on Pure and Applied Mathematics*, vol. 41, no. 7, pp. 909–996, OCT 1988.

[51] W. S. Lee and A. Kassim, "Signal and image approximation using interval wavelet transform," *IEEE Transactions on Image Processing*, vol. 16, no. 1, pp. 46–56, 2007.

[52] I. M. Johnstone and B. W. Silverman, "Boundary coiflets for wavelet shrinkage in function estimation," *Journal of Applied Probability*, vol. 41 A, no. SPEC. ISSUE, pp. 81–98, 2004.

[53] L. Brechet, M.-F. Lucas, C. Doncarli, and D. Farina, "Compression of biomedical signals with mother wavelet optimization and best-basis wavelet packet selection," *IEEE Transactions on Biomedical Engineering*, vol. 54, no. 12, pp. 2186–2192, DEC 2007.